Fisheries and Oceans Canada | Pêches et Océans Canada
Science | Sciences

**CSAS** | **SCCS**

**Canadian Science Advisory Secretariat** | **Secrétariat canadien de consultation scientifique**

**Proceedings Series 2006/042** | **Série des comptes rendus 2006/042**

# Proceedings of the Assessment Framework Workshop on Southern Gulf of St. Lawrence Snow Crab (Areas 12, E, F and 19) Gulf Regional Advisory Process

**11 – 14 October 2005**

**Miramichi Boardroom**
**Gulf Fisheries Centre**
**Moncton, New Brunswick**

**R. O'Boyle (Chair)**

# Compte rendu de l'atelier sur le cadre d'évaluation du crabe des neiges du sud du golfe du Saint-Laurent (zones 12, E, F et 19) Processus consultatif régional du Golfe

**Du 11 au 14 octobre 2005**

**Salle de conférence Miramichi**
**Centre des pêches du Golfe**
**Moncton (Nouveau-Brunswick)**

**R. O'Boyle (président)**

Bedford Institute of Oceanography / Institut océanographique de Bedford
1 Challenger Drive / 1, promenade Challenger
P.O. Box 1006 / C.P. 1006
Dartmouth, Nova Scotia / Dartmouth (N.-É.)
B2Y 4A2

**May 2007** | **mai 2007**

Canada

## FOREWORD

These Proceedings are a record of RAP meetings which were held during 11–14 October 2005. The report records as faithfully as possible the contributions and discussion that transpired at the meetings. However, the individual interpretations and opinions expressed at the meeting are not necessarily or in all cases scientifically sustainable or supported by other participants. The discussion summaries document the deliberations, which led to the tabled proposals. No statements are to be taken as reflecting the consensus of the meeting unless they are clearly identified as such. Moreover, additional information and further review may result in a change of decision where tentative agreement has been reached.

## AVANT – PROPOS

Le présent compte rendu relate les travaux des réunions tenues dans le cadre du PCR du 11 au 14 octobre 2005. Il reflète aussi fidèlement que possible les contributions et commentaires des participants à la réunion. Toutefois, les opinions et interprétations individuelles qui y sont présentées ne sont pas nécessairement ou toujours soutenables sur le plan scientifique, ou appuyées par les autres participants. Le résumé des discussions documente les délibérations ayant abouti aux propositions déposées. Aucune déclaration ne doit être considérée comme une expression du consensus des participants, sauf s'il est clairement indiqué qu'elle l'est effectivement. En outre, des renseignements supplémentaires et un plus ample examen peuvent avoir pour effet de modifier une décision qui avait fait l'objet d'un accord préliminaire.

2008-02810/



# Proceedings of the Assessment Framework Workshop on Southern Gulf of St. Lawrence Snow Crab (Areas 12, E, F and 19) Gulf Regional Advisory Process

**11 – 14 October 2005**

**Miramichi Boardroom**
**Gulf Fisheries Centre**
**Moncton, New Brunswick**

**R. O'Boyle (Chair)**

# Compte rendu de l'atelier sur le cadre d'évaluation du crabe des neiges du sud du golfe du Saint-Laurent (zones 12, E, F et 19) Processus consultatif régional du Golfe

**Du 11 au 14 octobre 2005**

**Salle de conférence Miramichi**
**Centre des pêches du Golfe**
**Moncton (Nouveau-Brunswick)**

**R. O'Boyle (président)**

Bedford Institute of Oceanography / Institut océanographique de Bedford
1 Challenger Drive / 1, promenade Challenger
P.O. Box 1006 / C.P. 1006
Dartmouth, Nova Scotia / Dartmouth (N.-É.)
B2Y 4A2

**May 2007**

**mai 2007**

© Her Majesty the Queen in Right of Canada, 2006
© Sa Majesté la Reine du chef du Canada, 2006

ISSN 1701-1272 (Printed / Imprimé)

Published and available free from:
Une publication gratuite de :

Fisheries and Oceans Canada / Pêches et Océans Canada
Canadian Science Advisory Secretariat / Secrétariat canadien de consultation scientifique
200, rue Kent Street
Ottawa, Ontario
K1A 0E6

http://www.dfo-mpo.gc.ca/csas/

CSAS@DFO-MPO.GC.CA

Printed on recycled paper.
Imprimé sur du papier recyclé.

Correct citation for this publication:

DFO, 2006. Proceedings of the Assessment Framework Workshop on Southern Gulf of St. Lawrence Snow Crab (Areas 12, E, F and 19), Gulf Regional Advisory Process; 11-14 October 2005. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2006/042.

On doit citer cette publication comme suit :

MPO, 2006. Compte rendu de l'atelier sur le cadre d'évaluation du crabe des neiges du sud du golfe du Saint-Laurent (zones 12, E, F et 19), Processus consultatif régional du Golfe; Du 11 au 14 octobre 2005. Secr. can. de consult. sci. du MPO, Compte rendu. 2006/042.

# TABLE OF CONTENTS / TABLE DES MATIÈRES

ABSTRACT / RÉSUMÉ ..... iv

INTRODUCTION / INTRODUCTION ..... 1

PRESENTATIONS / EXPOSÉS ..... 3

Wade, E., T. Surette, J. Marcil, M. Moriyasu, M. Hébert et M. Biron. 2005. Methodology in the Estimation of Snow Crab Resource in the Southern Gulf of St. Lawrence, RAP Working Paper 2005/12 / Méthode d'estimation de la ressource que constitue le crabe des neiges dans le sud du golfe du Saint Laurent. Document de travail du PCR 2005/12 ..... 3

Conan, G. Geostatistical Techniques and Examples / Techniques géostatistiques et exemples ..... 8

Choi, J. Survey Treatment of 4VW Snow Crab / Traitement des données du relevé du crabe des neiges de 4VW ..... 10

SYNOPSIS OF ISSUES / SYNTHÈSE DES POINTS ..... 12

Spatial and Temporal Distribution of Sampling / Distribution spatiale et temporelle de l'échantillonnage ..... 12

Kriging Algorithm / Algorithme du krigeage ..... 14

Pseudo Zeros / Pseudozéros ..... 16

Geographically Delimited Estimates of Biomass / Estimations délimitées géographiquement de la biomasse ..... 16

RESOLUTION OF ISSUES / RÈGLEMENT DES POINTS ..... 16

Impact of Spatial and Temporal Heterogeneity in Sampling Distribution / Répercussions de l'hétérogénéité spatiale et temporelle sur la distribution d'échantillonnage ..... 16

Kriging Algorithm and Spatial Boundaries of Snow Crab Habitat / Algorithme du krigeage et limites spatiales de l'habitat du crabe des neiges ..... 18

Pseudo Zeros / Pseudozéros ..... 20

Procedures to Produce Geographically Delimited Estimate of Biomass / Procédures à suivre pour produire une estimation géographiquement délimitée de la biomasse ..... 20

Additional Issue on Sex Ratios in 2005 Survey / Point additionnel sur les rapports des sexes dans le relevé de 2004 ..... 21

CONCLUDING REMARKS / MOT DE LA FIN ..... 21

APPENDICES / APPENDICES ..... 22

Appendix 1. List of Participants / Appendice 1. Liste des participants ..... 22
Appendix 2. Remit / Appendice 2. Mandat ..... 23
Appendix 3. Agenda / Appendice 3. Ordre du jour ..... 26
Appendix 4. Exploratory Analyses Undertaken at the Meeting / Appendice 4. Analyses exploratoires entreprises à la réunion ..... 28

## ABSTRACT

The assessment of the Gulf of St. Lawrence snow crab resource has been based upon geospatial statistical analyses of the summer survey information for a number of years. This approach has also been used when responding to management questions that have arisen from time to time. A number of issues have been raised in regards to the technical details of not only the post-survey treatment of the data but also of the survey design (e.g. boundary conditions). The current meeting afforded an opportunity for experts on these topics to discuss the issues and recommend both shorter-term solutions and longer – term research activities.

## RÉSUMÉ

L'évaluation de la ressource que constitue le crabe des neiges du golfe du Saint-Laurent est fondée depuis un certain nombre d'années sur des analyses statistiques géospatiales des données du relevé d'été. Cette méthode a aussi été utilisée pour répondre aux questions de gestion qui se sont posées de temps à autre. On a soulevé un certain nombre de questions au sujet des détails techniques non seulement du traitement après le relevé des données, mais également à propos du plan du relevé (des conditions à la limite, notamment). La réunion a offert une occasion à des experts de ces questions de discuter des points s'y rattachant et de recommander des solutions à court terme et des activités de recherche à long terme.

## INTRODUCTION

After welcoming the participants (Appendix 1) and introductions around the room, the chair (R. O'Boyle) reviewed the background of the meeting. He pointed out that since 2001, the Gulf and Maritimes Regional Advisory Processes have restructured the review meetings so that the technical review of the frameworks to be used in the annual assessments were undertaken separately from the annual assessment meetings. This is to allow more in-depth review of the technical basis of the assessments. The annual assessment meetings were to focus upon quality assurance of application of the framework as well as drafting of the Canadian Science Advisory Secretariat Science Advisory Report. For Gulf of St. Lawrence snow crab, the 2005 assessment meeting was held during 22 – 24 February 2005 and the 2006 assessment was planned for 28 February – 2 March 2006.

Framework reviews can be undertaken on all or part of the following:

a. Definition of management unit
b. Best approach to determine stock status

c. Reference points
d. Projection procedures

The current meeting is a framework review of the treatment of the 4T snow crab survey data, specifically (see Appendix 2 for more details):

- Exploration of the impact of heterogeneity in the spatial and temporal distribution of sampling sites on the biomass estimates and variograms.

- Investigation of the kriging algorithm, particularly in regards to establishing spatial boundaries of snow crab habitat.

- In cases where pseudo zeros are used, considerations are made on their number and location.

## INTRODUCTION

À la suite du mot de bienvenue aux participants (appendice 1) et des tours de table et présentations des personnes présentes, le président (R. O'Boyle) situe l'objet de la réunion dans son contexte. Il fait observer que depuis 2001, les processus consultatifs régionaux du Golfe et des Maritimes ayant entraîné la restructuration des réunions d'examen, on a entrepris l'examen technique du cadre à utiliser à l'occasion des évaluations annuelles séparément des réunions annuelles d'évaluation. Cela vise à permettre d'effectuer un examen plus approfondi du fondement technique des évaluations. Les réunions annuelles d'évaluation devaient être axées sur l'assurance de la qualité d'application du cadre et sur la rédaction d'une ébauche du rapport consultatif scientifique du Secrétariat canadien de consultation scientifique. Pour le crabe des neiges du golfe du Saint-Laurent, la réunion d'évaluation de 2005 a eu lieu du 22 au 24 février 2005 et l'évaluation de 2006 étaitt prévue pour la période comprise entre le 28 février et le 2 mars 2006.

On peut entreprendre des examens du cadre d'évaluation en se fondant sur la totalité ou sur une partie de ce qui suit :

a. La définition d'une unité de gestion
b. La meilleure méthode pour déterminer l'état d'un stock
c. Les points de référence
d. Les procédures de projection

La réunion a pour objet un examen du cadre d'évaluation du traitement des données du relevé du crabe des neiges de 4T, portant tout spécialement sur (voir l'appendice 2 pour plus de détails) :

- L'analyse des répercussions de l'hétérogénéité à l'intérieur de la distribution spatiale et temporelle de sites d'échantillonnage sur les estimations et les variogrammes de la biomasse.

- L'étude de l'algorithme du krigeage, notamment en ce qui concerne l'établissement de limites spatiales de l'habitat du crabe des neiges.

- Dans les cas où on utilise des pseudozéros, l'examen de leur nombre et de leurs emplacements.

- Investigation of procedures using a subset of the survey data to produce geographically delimited estimate of biomass for management purposes

The chair then reviewed the agenda (Appendix 3), emphasizing that this was a guide to the presentation and discussion and would be open to modification according to the needs of the meeting. Time was available in the agenda to undertake exploratory analyses in support of the remit.

M. Lanteigne was assigned as the meeting rapporteur. As well, the discussion was recorded on tape to facilitate post-meeting compilation of the proceedings.

After questions of clarification, the presentations commenced.

- L'étude de procédures faisant appel à un sous-ensemble des données du relevé pour produire une estimation géographiquement délimitée de la biomasse à des fins de gestion

Le président passe ensuite en revue l'ordre du jour (appendice 3), en soulignant qu'il s'agit d'un guide pour les exposés et les discussions et qu'il se montrera ouvert à des modifications selon les besoins des participants à la réunion. On dispose de temps à l'intérieur de l'ordre du jour pour entreprendre des analyses exploratoires destinées à appuyer le mandat.

On désigne M. Lanteigne rapporteur pour la réunion. On enregistre également les discussions sur bande pour faciliter la compilation après la réunion de son compte rendu.

Les exposés débutent après les questions d'éclaircissement.

## PRESENTATIONS

**Wade, E., T. Surette, J. Marcil, M. Moriyasu, M. Hebért, and M. Biron. 2005. Methodology in the Estimation of Snow Crab Resource in the Southern Gulf of St. Lawrence. RAP Working Paper 2005/12**

**Hebert, M., E. Wade, T. Surette, P. DeGrace, and R. Ruest. 2005. Stock Assessment, Fisheries, Life Cycle of Snow Crab (Chionoecetes opilio) in the southern Gulf of St. Larwence (Areas 12, 19, E and F) and Table / Figures RAP Working Paper 2005/11A and B**

***Presentation Highlights***

There are four issues involved in biomass index estimation, these relating to both sampling design and specific aspects of kriging, particularly the analytical method used.

1st Issue (Variable Sampling Density)

The trawl survey begun in 1988 was initially based on a stratified random design and has evolved over time. As the fishery developed, the survey area was expanded to accommodate new fishing grounds. Furthermore, some fishing areas were created while others were amalgamated. As a consequence, the number of sample points has varied through both space and time, as the size of the zone to be sampled has expanded and the local density of sample points has also varied. Subjective decisions were made to redistribute the samples from one zone to another based upon specific management concerns. These factors resulted in a non uniform distribution of the survey area. In order to compensate for this effect, a sub-sample built by taking a single sample from each design grid is used to fit the variogram.

2nd Issue: Two Variograms (Global and Local)

Two variograms are used to estimate abundance indices in the southern Gulf of St. Lawrence. The first is based upon the global southern Gulf of St. Lawrence (sGSL) dataset and is used to estimate biomass in areas 12 and E. The second

## EXPOSÉS

**Wade, E., T. Surette, J. Marcil, M. Moriyasu, M. Hébert et M. Biron. 2005, Méthode d'estimation de la ressource que constitue le crabe des neiges dans le sud du golfe du Saint-Laurent. Document de travail du PCR 2005/12**

**Hebert, M., E. Wade, T. Surette, P. DeGrace et R. Ruest. 2005, Évaluation du stock, pêches, cycle de vie du crabe des neiges (*Chionoecetes opilio*) dans le sud du golfe du Saint-Laurent (zones 12, 19, E et F) et tableau / figures – Documents de travail du CPR 2005/11A et B**

***Points saillants des exposés***

Il y a quatre points associés à l'estimation de l'indice de la biomasse; ces points se rattachent aussi bien au plan d'échantillonnage qu'à des aspects précis du krigeage, à la méthode d'analyse utilisée en particulier.

1er point (densité d'échantillonnage variable)

Le relevé au chalut qui a débuté en 1988 a au départ été fondé sur un plan aléatoire stratifié et a évolué au fil du temps. À mesure que la pêche s'est développée, on a élargi la zone du relevé pour y ajouter de nouveaux lieux de pêche. On a, en outre, créé de nouvelles zones de pêche, tandis qu'on en a amalgamé d'autres. Par conséquent, le nombre de points échantillons varie tant dans le temps que dans l'espace, étant donné que les dimensions de la zone à échantillonner ont augmenté; la densité locale des points échantillons varie également. Il a fallu prendre des décisions subjectives pour redistribuer les échantillons d'une zone à une autre en se fondant sur des préoccupations bien précises en matière de gestion. Ces facteurs ont entraîné une distribution non uniforme de la zone du relevé. Afin de compenser cet effet, on utilise un sous-échantillon construit en prélevant un seul échantillon de chaque grille d'échantillonnage pour ajuster le variogramme.

2e point : Deux variogrammes (global et local)

On utilise deux variogrammes pour estimer les indices d'abondance dans le sud du golfe du Saint-Laurent. Le premier est fondé sur tout l'ensemble de données du sud du golfe du Saint-Laurent (sGSL) et sert à estimer la

is fitted to data from the Cape-Breton subunit which comprises areas 19, F and parts of the eastern part of 12. Variogram parameters for area 12 were compared with population estimates. The range and sill parameters of the variograms were found to be correlated with the population estimates.

biomasse dans les zones 12 et E. Le second est ajusté aux données provenant de la sous-unité du Cap-Breton, qui comprend les zones 19 et F et des parties de la portion orientale de la zone 12. On a comparé les paramètres des variogrammes pour la zone 12 aux estimations de la population. On a constaté une corrélation entre les paramètres « zone d'influence » et « plateau » des variogrammes et les estimations de la population.

3rd Issue: Sector 1 (Small Subareas and Bias)

Sector one of Area 12 (Bay of Chaleur region) is a much contested fishing ground in the snow crab industry. Its proximity to local fishing ports makes it economically advantageous for fishermen to first exploit these areas before moving on to other grounds. In recent years, observed sex ratios based upon the snow crab survey have shown that fewer commercial males are available for mating in Sector one than for the rest of the sGSL fishing areas. The proportions of soft crab in the catches have been a major problem in the past in Sector one, with closures of parts or of the entire zone occurring regularly since the last decade. The high occurrence of soft crab in the catches is often linked to over-exploitation. Despite the importance of Sector one, a number of analytical problems prevent an accurate assessment of its population size. In addition, the precision of the estimate is limited by the small size of Sector one, and results in large confidence intervals. Also, simulations have shown that the areas sampled within were biased, especially during the early years, with shallow waters being less well covered than deeper waters.

3e point : Secteur 1 (petites sous-zones et biais)

Le secteur 1 de la zone 12 (la région de la baie des Chaleurs) est un lieu de pêche fortement contesté dans l'industrie du crabe des neiges. Sa proximité par rapport à des ports de pêche de la région rend économiquement avantageux pour les pêcheurs de l'exploiter d'abord avant de passer à d'autres lieux d'exploitation. Ces dernières années, les rapports des sexes observés à partir du relevé du crabe des neiges ont montré qu'il y a moins de mâles de taille commerciale disponibles pour la reproduction que dans le reste des zones de pêche du sGSL. Les proportions de crabes à carapace molle à l'intérieur des prises ont constitué un problème majeur dans le passé à l'intérieur du secteur 1, des fermetures de parties ou de la totalité de la zone se produisant régulièrement depuis la dernière décennie. La fréquence élevée des crabes à carapace molle à l'intérieur des prises est souvent liée à la surexploitation. Malgré l'importance du secteur 1, un certain nombre de problèmes analytiques empêche d'évaluer exactement la taille de la population de crabes. La précision de l'estimation est, en plus, limitée par les petites dimensions du secteur 1 et entraîne de gros intervalles de confiance. Des simulations ont également montré que les zones échantillonnées dans le secteur étaient biaisées, surtout les premières années, les eaux peu profondes étant moins bien recouvertes que les eaux plus profondes.

4th Issue: Pseudo Zeros

'Pseudo-zeros' have been used in the estimation of crab density and biomass to describe the limit of the snow crab habitat, by projecting the probable absence of snow crab population due to high water temperature where very shallow or very deep waters occur in the sGSL. Survey data from both the snow crab bottom trawl survey and groundfish September survey, along with specific experiments looking at crab distribution with depth, show that the population dissipates well

4e point : Pseudozéros

On utilise des « pseudozéros » à l'intérieur de l'estimation de la densité de crabes et de leur biomasse pour décrire les limites de l'habitat du crabe des neiges, en projetant l'absence probable de la population de ces crabes en raison de la température élevée de l'eau là où il y a dans le sGSL des eaux très peu profondes ou très profondes. Des données tirées à la fois du relevé au chalut de fond du crabe des neiges et du relevé de septembre du poisson de fond,

before approaching water depths of 36.5m and that at 20 meters, snow crab is virtually absent. Although the boundaries of snow crab distribution are certainly dynamic, the 20m depth used for kriging extends well beyond these boundaries. Using 'pseudo-zeros' in kriging has a varied effect on the population estimates. For the estimates of area 12, a difference of 5 to7% increase in population is calculated in recent years when 'pseudo-zeros' are omitted. In Sector one, the difference is much greater, ranging from 40 to 80% increase. The effect of pseudo-zeros is mostly at the edge of the snow crab survey areas such as Sector one, north of PEI, along the Magdalene Islands and along Area E. A transect simulation study conducted in Sector one showed that the global mean is sensitive to the number of pseudo-zeroes, though in practice its effect on kriging is influenced by the number of true sample points, the type of variogram, the size of the search neighborhood and the presence of a spatial sampling bias. Another alternative to the 'pseudo-zeros' would be to trawl beyond the 20m line. However, trawling in these shallow areas is not practical and the high cost associated with this may be difficult to justify.

parallèlement à des expériences précises consacrées à l'étude de la distribution du crabe suivant la profondeur, montrent que la population de crabes se disperse bien avant qu'on s'approche d'une profondeur d'eau de 36,5 m et qu'à 20 m il n'y en a pratiquement pas. Même si les limites de la distribution du crabe des neiges sont certainement dynamiques, la profondeur de 20 m utilisée pour le krigeage s'étend bien au-delà de ces limites. L'utilisation de « pseudozéros » pour le krigeage a des effets variables sur les estimations de la population. Pour les estimations relatives à la zone 12, on calcule qu'il y a eu une augmentation de la population de 5 à 7 % ces dernières années lorsqu'on omet les « pseudozéros ». Dans le secteur 1, la différence est beaucoup plus grande, allant d'une augmentation de 40 à 80 %. L'effet des pseudozéros se fait principalement sentir à la bordure des zones de relevé du crabe des neiges comme le secteur 1, la côte nord de l'Î.-P.-É., le littoral des Îles-de-la-Madeleine et celui de la zone E. Une étude en simulation de transects menée dans le secteur 1 a montré que la moyenne globale est sensible au nombre de pseudozéros, même si en pratique son effet sur le krigeage est influencé par le nombre de vrais points échantillons, le type de variogramme, les dimensions du voisinage des recherches et la présence d'un biais d'échantillonnage dans l'espace. Une autre solution aux « pseudozéros » consisterait à effectuer du chalutage au-delà de la ligne des 20 m. Le chalutage dans ces zones aux eaux peu profondes n'est cependant pas pratique et le coût élevé qui y est associé peut être difficile à justifier.

***Discussion***

- Explanation was requested on the sampling allocation, specifically, the distribution of samples as the number of stations increased. It was shown that while most grids in area 12 contained one sample, many contained two or three samples. In the Cape Breton (CB) sub-unit specifically, most grids have two or three samples.

- Explanation was also requested on the number of samples used in modeling the variogram for the CB sub-unit; it was shown that in the past, all samples were used from the CB sub-unit to produce the CB sub-unit variogram. A local variogram was used since it was believed that some characteristics of

***Discussion***

- On demande des explications au sujet du plan d'échantillonnage, en particulier à propos de la distribution des échantillons, parce que le nombre de stations a augmenté. On montre que même si la plupart des grilles dans la zone 12 contenaient un seul échantillon, beaucoup en renfermaient deux ou trois. Dans la sous-unité du Cap-Breton (C-B), notamment, la plupart des grilles contenaient deux ou trois échantillons.

- On demande aussi des explications au sujet du nombre d'échantillons utilisés dans le cadre de la modélisation du variogramme pour la sous-unité du C-B; on indique que dans le passé, on a utilisé tous les échantillons de la sous-unité du C-B pour produire le variogramme de cette dernière.

the variogram in this local area were different than for the entire southern Gulf of St. Lawrence area.

On a utilisé un variogramme local, parce qu'on croyait que certaines caractéristiques du variogramme dans cette zone étaient différentes de celles de l'ensemble de la zone sud du golfe du Saint-Laurent.

- Some participants suggested that it might be more robust to only use the variograms to produce maps and not be used for variance estimation.

- Certains participants laissent entendre qu'il serait peut-être plus robuste statistiquement de n'utiliser les variogrammes que pour produire des cartes et de ne pas les employer pour l'estimation de la variance.

- There was some discussion on the definitions of stratified *vs* lattice sampling. The present scheme used in the snow crab survey could be considered as lattice sampling.

- On discute un peu de la définition de l'échantillonnage stratifié par opposition à celle de l'échantillonnage en treillis ou en grille. On pourrait considérer comme de l'échantillonnage en treillis le plan actuel utilisé dans le cadre du relevé du crabe des neiges.

- Discussion then turned to the question of how we should respond in the future to sampling design changes or grid reallocation requests. One participant suggested that an answer may reside outside traditional geostatistics and provided input on the subject of attaining more robust estimates from small areas. Suggestions included sampling globally to meet local need, using known small area estimation techniques and GLM using random effects. The participant also stated that he had some ideas on how to pursue this during this workshop.

- On discute ensuite de la question consistant à savoir comment nous devrions à l'avenir réagir aux changements au plan d'échantillonnage ou aux demandes de réattribution des grilles. Un participant laisse entendre que l'une des réponses à la question peut provenir d'ailleurs que de la géostatistique traditionnelle et formule des commentaires sur la réalisation d'estimations plus robustes à partir de petites zones. Il suggère notamment un échantillonnage global pour répondre à un besoin au niveau local, l'utilisation de techniques d'estimation dans de petites zones connues et un MLG faisant appel à des effets aléatoires. Le participant déclare également qu'il a quelques idées sur la façon de poursuivre l'étude de cette question durant l'atelier.

- Questions were asked on the variogram construction. With approximately 45 categories of crabs, variograms are produced each year for each category. These variograms may be different from each other. However, the main variograms used are those for the commercial males, females, R1, R2, R3 and R4 crab categories.

- On pose des questions au sujet de la construction des variogrammes. Étant donné qu'il existe environ 45 catégories de crabes, on produit chaque année des variogrammes pour chaque catégorie. Ces variogrammes peuvent être différents les uns des autres. Les principaux variogrammes utilisés sont cependant ceux associés aux crabes mâles de taille commerciale, de crabes femelles et aux crabes R1, R2, R3 et R4.

- A participant requested details on how the 45 categories are being used and selected in the assessment. The presenter responded that all these categories group crabs were based on sex, shell condition, maturity and sizes.

- Un participant demande des détails sur la façon dont on utilise et on sélectionne actuellement les 45 catégories dans le cadre de l'évaluation. Le présentateur répond que toutes ces catégories de crabes reposent sur le sexe, l'état de la carapace, la maturité et la taille.

- A participant indicated that the definition of each category was important and must be clearly specified. He also argued that it was possible that his definition of some of the female categories (mature, primiparous and pre primiparous for example) might be different than what was used in the present assessment. The question was asked on how the verification of those categories, especially for the females, was conducted. The presenter explained the definition and nomenclature used for the female categories. Similarly, there was discussion on the definition of skip moulters and their implication in the assessment process. The presenter indicated that this specific issue of category identification is more a biological issue. It is an interesting topic to discuss but was outside the remit of this framework meeting and would have no impact on the analytical approach being discussed. The participant's disagreement was noted.

- Un participant mentionne que la définition de chaque catégorie est importante et qu'il faut clairement la préciser. Il soutient aussi qu'il est possible que sa définition à lui de certaines des catégories de femelles (matures, primipares et pré-primipares, par exemple) pourrait être différente de celle utilisée dans le cadre de l'évaluation actuelle. On demande comment on vérifie ces catégories, les catégories de femelles principalement. Le présentateur explique la définition et la nomenclature utilisées pour les catégories de femelles. On discute également de la définition des individus ayant sauté une mue et de leurs conséquences à l'intérieur du processus d'évaluation. Le présentateur mentionne que cette question précise de l'identification des catégories est davantage une question d'ordre biologique. C'est un sujet intéressant à traiter, mais qui ne s'inscrit pas à l'intérieur du mandat des participants à la réunion sur le cadre d'évaluation et qui n'aura aucune répercussion sur la méthode d'analyse dont on est en train de discuter. On note que le participant susmentionné n'est pas d'accord.

- It was requested that variograms be further studied to understand how they vary over time. It was shown that they varied according to their spatial spread and abundance. Also, the sill of the variogram model varies according to the population size.

- On demande d'étudier plus à fond les variogrammes pour comprendre en quoi ils varient au fil du temps. On a constaté qu'ils varient suivant l'étalement et l'abondance du crabe dans l'espace. Le plateau du modèle de variogramme varie également suivant la taille de la population.

- Discussions and questions then centered on the definition of the area used in population estimation. Is this area delineation set in stone? Which came first, the delineation of the area or survey design? It was explained that it was originally based on fishermen's information and logbooks back in the mid 1980s (page 59 of Hebért et. al., 2005 shows fishing distributions). Therefore, the areas came first and then the design followed.

- La discussion et les questions sont ensuite centrées sur la définition de la zone utilisée dans le cadre de l'estimation de la population. Cette délimitation de la zone est-elle coulée dans le béton? Qu'est-ce qui prime? La délimitation de la zone ou le plan d'échantillonnage? On explique que la délimitation était au départ fondée sur des données des pêcheurs et des registres de bord remontant au milieu des années 80 (la page 59 de Hébert et coll., 2005 illustre la répartition des lieux de pêche). Les zones priment, par conséquent, sur le plan.

**Conan, G. Geostatistical Techniques and Examples**

***Presentation Highlights***

Presentation highlights not available

***Discussion***

- There was discussion on the survey gear type and gear mensuration (measurement of gear during trawling). Explanations were given on the types of measurements presently used in the Gulf survey and the filtering process used. Although it was considered an interesting issue, it was determined that the workshop was not the venue to sort out the details of gear measurement techniques and filtering of these data. Other participant pointed out that these errors in measurement would be taken into account in the nugget effect of the variogram. There were differences of opinion on this issue.

- A participant pointed that three vessel changes had occurred over time in the Gulf survey and that this factor alone is perhaps more important than the issue of gear measurements.

- It was pointed out that survey cell size is not constant over the Gulf area. A 10 minute by 10 minute grid is larger in the southern latitudes than in the north, and that the presenter's software corrects for this by creating similar sized blocks for establishing the sampling design. Therefore, no bias is created between samples placed in the north or south. Actual measurements of grid areas in the Gulf showed 3% difference in surface between northern and southern grids. This factor was not considered significant by other participants.

- The presenter also pointed out that when converting from latitude-longitude coordinates to flat x-y distance, one needs to account for the curvature of earth. The Gulf software uses a UTM projection used to convert coordinates and was deemed to account for this factor.

**Conan, G. Techniques géostatisques et exemples**

***Points saillants de l'exposé***

Les points saillants de l'exposé ne sont pas disponibles.

***Discussion***

- On discute du type d'engin utilisé pour le relevé et des dimensions de cet engin (de ses mesures durant le chalutage). On fournit des explications sur les types de mesures actuellement utilisés dans le cadre du relevé effectué dans le golfe et du procédé de filtrage employé. Même si l'on considère qu'il s'agit d'un point intéressant, on détermine que l'atelier n'est pas le lieu où régler les détails des techniques de mesure de l'engin et du filtrage de ces données. Un autre participant fait remarquer qu'on tiendra compte de ces erreurs de mesure à l'intérieur de l'effet de pépite du variogramme. Il y a des divergences de vues sur ce point.

- Un participant mentionne qu'on a changé trois fois de bateau au fil du temps dans le cadre du relevé effectué dans le golfe et que ce facteur à lui seul est peut-être plus important que la question des mesures de l'engin.

- On fait remarquer que les dimensions des mailles de la grille du relevé ne sont pas constantes au-dessus de la zone du golfe. Une grille de 10 minutes x 10 minutes est plus grande aux latitudes méridionales qu'aux latitudes septentrionales, ce que le logiciel du présentateur corrige en créant des blocs de dimensions similaires pour l'établissement du plan d'échantillonnage. Il n'y a donc pas création de biais entre les échantillons placés au nord ou au sud. Les mesures réelles des zones quadrillées dans le golfe ont montré une différence de 3 % à la surface entre les grilles du nord et du sud. D'autres participants ne considèrent pas ce facteur significatif.

- Le présentateur fait aussi remarquer qu'on doit tenir compte de la sphéricité de la Terre quand on convertit des coordonnées de latitude et de longitude en distance x-y plane. Le logiciel utilisé dans la région du golfe fait appel à une projection de Mercator transverse universel (MTU) qu'on emploie pour convertir des coordonnées; il a, en

- Discussion then focused on the importance of the definition of the field and estimation method.

- Regarding variogram fitting, the presenter stated that he fits all variograms manually/visually and indicates that discretion of the user is important.

- Discussion then centered on the importance of the level of the sill of the variogram and the importance of determining its shape. The decision on whether to use an exponential, spherical, linear or Gaussian model will have repercussions on the results.

- It was considered that the variograms presently being used are adequate.

- There was discussion on whether it was worthwhile to review crab categories, especially for the commercial male categories and mature female categories.

- The use of pseudo-zeros was then discussed. An alternative to pseudo zeros was also introduced by a participant. The presenter explained that kriging with external drift produces the same goal as one with pseudo zeros by taking into account the depth *vs.* density relationship

- Calculation methods were also revised.

- Lastly, discussions centered of the differences that may occur between the results from the simple mean calculations *vs.* the kriging estimates. They should not be very different when the same points are used. Factors such as sparse datasets, sampling density, edge effects and others can cause discrepancies between the two methods of calculation.

outre, la réputation de tenir compte de ce facteur.

- La discussion est ensuite axée sur l'importance de la définition du champ et de la méthode d'estimation.

- Au sujet de l'ajustement des variogrammes, le présentateur déclare qu'il ajuste tous les variogrammes manuellement/visuellement et mentionne que le jugement de l'utilisateur est important.

- La discussion est ensuite centrée sur l'importance du niveau du plateau du variogramme et sur l'importance également d'en déterminer la forme. La décision d'utiliser un modèle exponentiel, sphérique, linéaire ou normal (gaussien) aura des répercussions sur les résultats.

- On considère que les variogrammes actuellement utilisés sont adéquats.

- On discute de la question consistant à savoir s'il vaut la peine d'examiner les catégories de crabes, surtout les catégories de mâles de taille commerciale et les catégories de femelles matures.

- Puis on discute de l'utilisation des pseudozéros. Un participant présente aussi une solution de rechange aux pseudozéros. Le présentateur explique que le krigeage avec dérive externe (KDE) produit le même objectif que le krigeage avec des pseudozéros en tenant compte de la relation entre la profondeur et la densité.

- On révise également les méthodes de calcul.

- La discussion tourne finalement autour des différences qui peuvent surgir entre les résultats découlant des calculs de moyennes simples par rapport aux estimations par krigeage. Les résultats ne devraient pas être très différents lorsqu'on utilise les mêmes points. Des facteurs comme des ensembles de données clairsemées, la densité d'échantillonnage, les effets de bord et d'autres facteurs risquent d'entraîner des divergences entre les deux méthodes de calcul.

**Choi, J. Survey Treatment of 4VW Snow Crab**

***Presentation Highlights***

The approaches used by the 4VWX snow crab assessments (DFO Science Maritimes Region) were summarized as they relate to the issues raised for consideration during the workshop. The Maritimes' contributions were relevant to objectives 2 and 4 of the Remit.

Objective 2: Snow Crab Habitat

The determination of the spatial bounds of snow crab habitat was addressed in the Maritimes Region by examining the historical records of snow crab distributions as a function of depth and bottom temperature at the time of sampling. The bounds of these distributions were used to identify a dynamic surface where snow crabs are likely to be found. Other higher-order factors (first and second derivatives) such as bottom slope, bottom curvature and bottom temperature gradients have been explored and were found to add some additional limitations to the bounds of potential snow crab habitat. However, first-order effects were noted as being the most influential in the determination of spatial bounds.

Objective 4: Biomass Estimation in Subareas

The estimation of biomass in geographical subareas was addressed by examining the consequences of forming a global (shelf-wide) variogram vs a local (sub-area) variogram. The forms of the empirical variograms were found to be quite variable and problems of model variogram stability and model convergence were observed. A simple solution of a three - year moving average was used to stabilize the variogram modeling. However, abundance estimates (from polygonal kriging predictions) for the sub-area were found to be quite stable, regardless of the variability between global vs local variograms. As a caveat, polygonal block kriging solutions were found to be stable only if the discretisation of solution blocks was at length scales of 1 to 5 km. Other approaches to the temporal stabilizations of variograms were explored during the workshop using various

**Choi, J. Traitement des données du relevé du crabe des neiges de 4VW**

***Points saillants de l'exposé***

On résume les méthodes utilisées dans le cadre des évaluations du crabe des neiges de VWX (MPO-Sciences, Région des Maritimes) qui se rattachent aux points soulevés à des fins d'examen durant l'atelier. La contribution de la Région des Maritimes est pertinente pour atteindre les objectifs 2 et 4 du mandat.

Objectif n° 2 : Habitat du crabe des neiges

On a abordé dans la Région des Maritimes la question de la détermination des limites spatiales de l'habitat du crabe des neiges en examinant les registres chronologiques des distributions de cette espèce en fonction de la profondeur et de la température au fond au moment de l'échantillonnage. On a utilisé les limites de ces distributions pour identifier une surface dynamique où l'on est susceptible de trouver des crabes des neiges. On a analysé d'autres facteurs d'ordre plus élevé (les premières et secondes dérivées) comme la pente du fond, la sphéricité de ce dernier et les gradients de température au fond et constaté qu'ils ajoutent certaines limites à celles de l'habitat potentiel du crabe des neiges. On a cependant noté que les effets du premier ordre sont ceux qui ont le plus d'influence sur la détermination des limites spatiales.

Objectif n° 4 : Estimation de la biomasse dans des sous-zones

On aborde la question de l'estimation de la biomasse dans des sous-zones géographiques en examinant les conséquences de la formation d'un variogramme global (englobant toute la plate-forme) par rapport à un variogramme local (de sous-zone). On a constaté que les formes des variogrammes empiriques varient assez et on a observé des problèmes de stabilité des variogrammes des modèles et de convergence de ces derniers. On a utilisé la solution simple d'une moyenne mobile de trois ans pour stabiliser la modélisation des variogrammes. On a cependant constaté que les estimations de l'abondance (établies à partir de prévisions par krigeage polygonales) pour la sous-zone étaient assez stables, indépendamment de la variabilité entre les variogrammes global et local. À titre d'avertissement, mentionnons qu'on a constaté que les solutions en matière de krigeage par

re-weighting and windowing schemes (1 to 10 yr moving windows and uniform/linear/exponential decay schemes). A three-year moving window was observed to provide sufficient stabilization for the 4T historical data.

***Discussion***

- A participant suggested to model variograms by year and then average over three years as this is the more standard methodology
- A rescaling methodology was also proposed. In his particular application, the technique suggested produced favorable results.
- It was pointed out that in the Gulf Region, as in Eastern Nova Scotia, crabs are attracted by pockets of cold water. Therefore, temperature is an important criterion in determining the area covering the stock.
- A participant pointed out that the size of the blocks can cause problems when performing kriging over an irregular shaped polygon. Large blocks can create boundary problems by not respecting the contour of the polygon while small blocks will contain little data. He suggested using an alternative numerical approach.
- According to a participant, block size used during block kriging used by the DFO Gulf Science group is not bad. The alternative otherwise is to go directly to polygon shaped block (numerical approach).
- The theory and application of co-kriging was discussed. The technique uses external variables (secondary variables) such as depth and or temperature to give shape to the kriging map. These secondary variables should be sampled more intensely than the primary variable (crab density) and should have an adequate link with the primary

blocs polygonales n'étaient stables que si la discrétisation des blocs de solution se situait à des échelles de longueur de 1 à 5 km. On analyse durant l'atelier d'autres approches pour les stabilisations temporelles des variogrammes faisant appel à divers plans de repondération et de fenêtrage (des fenêtres mobiles de 1 à 10 ans et des diagrammes de décroissance uniforme/linéaire/exponentiel). On a observé une fenêtre mobile de trois ans pour assurer une stabilisation suffisante des données historiques relatives à 4T.

***Discussion***

- Un participant suggère de modéliser les variogrammes par année et ensuite d'établir une moyenne sur trois ans, étant donné qu'il s'agit là de la méthode la plus courante.
- Il propose également une méthode de rééchelonnage. La technique suggérée a donné des résultats favorables dans le cadre de l'utilisation particulière qu'il en a faite.
- On fait remarquer que dans la Région du Golfe, comme à l'est de la Nouvelle-Écosse, les crabes sont attirés par les poches d'eau froide. La température est donc un critère important pour définir la zone englobant le stock.
- Un participant fait observer que la taille des blocs peut causer des problèmes lorsqu'on effectue du krigeage au-dessus d'un polygone de forme irrégulière. Les gros blocs peuvent créer des problèmes aux limites en ne respectant pas le contour du polygone, tandis que les petits renfermeront peu de données. Il suggère d'utiliser une méthode numérique de rechange.
- D'après un participant, la taille des blocs utilisée durant le krigeage par le groupe des Sciences de la Région du Golfe du MPO ne pose pas problème. L'alternative consiste à passer directement au bloc en forme de polygone (à la méthode numérique).
- On discute de la théorie et de l'application du co-krigeage. La technique fait appel à des variables externes (secondaires) comme la profondeur et/ou la température pour donner une forme à la carte de krigeage. On devrait échantillonner plus intensément ces variables secondaires qu'on ne le fait pour la variable primaire (la densité de crabes) et les relier

variable.

- It was also mentioned that co-kriging was attempted in the Gulf approximately seven years ago but results were not conclusive. Reasons for this may range from lack of data, (only eight years of data at the time) to lack of a relationship between temperature, depth and crab density.

adéquatement à la variable primaire.

- On mentionne également qu'on a tenté d'effectuer du co-krigeage dans le golfe il y a environ sept ans, mais que les résultats n'ont pas été concluants. Cela peut s'expliquer de diverses façons allant d'un manque de données (seulement huit ans de données à l'époque) à l'absence de relation entre la température, la profondeur et la densité de crabes.

## SYNOPSIS OF ISSUES

During the presentations, a number of issues were raised that were pursued through further analyses. These are outlined below according to the four objectives provided in the remit. The details of the analyses are provided in Appendix 4.

### Spatial and Temporal Distribution of Sampling

To introduce this topic, the chair posed a number of questions on the survey to stimulate the discussion. For instance,

- What is the primary objective of the sampling? To produce relative and / or absolute estimates of abundance or to produce distribution maps of biomass?

- What would be appropriate survey design and methodology for each objective?

During the subsequent discussion on these and other questions, it was felt that the existing design was adequate to address the questions posed of it but exploration of the construction of the variogram was required.

#### *Use of Past Survey Information*

The group discussed ways to make use of the historical information. Currently, each year's survey data are used independently to construct a variogram for that year. State - space modeling was mentioned as one possibility of making more use of the historical information. In this, under a Bayesian formulation, the posterior of one year could be used as the prior for the following year. To examine the potential impact of procedures

## SYNTHÈSE DES POINTS

On a soulevé durant les exposés un certain nombre de points dont on poursuit l'étude au moyen d'analyses plus poussées. Ces points sont présentés dans leurs grandes lignes ci-dessous suivant les quatre objectifs formulés dans le mandat. L'appendice 4 fournit les détails des analyses.

### Distribution spatiale et temporelle de l'échantillonnage

Pour présenter ce sujet, le président pose un certain nombre de questions sur le relevé afin de stimuler la discussion. Par exemple,

- Quel est le principal objectif de l'échantillonnage? Produire des estimations relatives et/ou absolues de l'abondance ou produire des cartes de distribution de la biomasse?

- Qu'est-ce qui constituerait un plan et une méthode de relevé appropriés pour chaque objectif?

Durant la discussion qui suit sur ces questions et d'autres, on estime que le plan actuel est adéquat pour aborder les questions posées à ce propos, mais qu'il faut étudier à fond la construction du variogramme.

#### *Utilisation de données de relevé passées*

Le groupe discute de moyens d'utiliser les données historiques. À l'heure actuelle, on emploie indépendamment les données du relevé de chaque année afin de construire un variogramme pour cette dernière. On mentionne que les modèles espace - état peuvent permettre d'utiliser davantage les données historiques. On pourrait dans le cadre de cette modélisation employer, aux termes d'une formulation de

that use information across a number of years, the performance of temporal moving averages in the stabilization of variograms was explored (Appendix 4A).

***Relationship amongst the Nugget, Range of Influence and Sill***

Discussion on the use of the historical information led to consideration on how the nugget, range of influence and sill might change with abundance (Appendix 4B) and investigation of the relationship between the slope at the origin and population size (Appendix 4C).

An analysis of how variograms changed over time (Appendix 4D) was instructive and suggested that one global variogram, scaled by the annual variance, across years was appropriate.

The issue of anisotropy was also discussed. This refers to when the spatial gradient in variance has a directional component e.g. variograms calculated using data selected from different quadrants around a central point will be different. An analysis using the survey data from Western Cape Breton (Appendix 4E) suggested that while anisotropy might be expected, it did not appear to be strongly supported by the data.

***Form of the Relationship***

The group then discussed whether or not there was an expected form to the variogram based upon theory. Is there a preferred model to use? It was suggested that a linear model used in statistical modeling could be used to investigate this (Appendix 4F) which could also be used to investigate the role of environmental parameters in influencing the variogram.

A number of explorations were undertaken to examine the sensitivity of the kriging to the parameters of the variogram (nugget, range of influence and sill). Overall, these explorations

Bayes, l'a posteriori d'une année en tant qu'a priori de l'année suivante. On analyse le rendement de moyennes mobiles temporelles dans le cadre de la stabilisation des variogrammes pour examiner les répercussions éventuelles de procédures faisant appel à des données recoupant un certain nombre d'années (appendice 4A).

***Relations entre la pépite, la zone d'influence et le plateau***

La discussion sur l'utilisation des données historiques amène à examiner en quoi la pépite, la zone d'influence et le plateau pourraient varier suivant l'abondance (appendice 4B) et à étudier la relation entre la pente à l'origine et la taille de la population (appendice 4C).

Une analyse de la nature des fluctuations des variogrammes au fil du temps (appendice 4D) est instructive et laisse supposer qu'un seul variogramme global, échelonné suivant la variance annuelle, d'une année à une autre, est approprié.

On discute aussi de la question de l'anisotropie. L'anisotropie renvoie au moment où le gradient spatial à l'intérieur de la variance a une composante directionnelle; des variogrammes calculés à l'aide de données sélectionnées à partir de différents quadrants autour d'un point central seront, par exemple, différents. Une analyse faisant appel aux données de relevé provenant de l'ouest du Cap-Breton (appendice 4E) laisse supposer qu'on pourrait s'attendre à de l'anisotropie, ce que les données ne semblent pas appuyer solidement cependant.

***Forme de la relation***

Le groupe discute ensuite de la question consistant à savoir si d'après la théorie, l'on s'attend à ce que le variogramme ait une forme quelconque. Y a-t-il un modèle qu'il est préférable d'utiliser? On laisse entendre qu'on pourrait employer un modèle linéaire utilisé en modélisation statistique pour étudier cette question (appendice 4F) et dont on pourrait aussi se servir pour étudier l'influence de paramètres environnementaux sur le variogramme.

On entreprend un certain nombre d'analyses pour examiner la sensibilité du krigeage aux paramètres du variogramme (la pépite, la zone d'influence et le plateau). Globalement, ces

showed that the kriging is relatively robust to changes in these parameters (Appendix 4G).

Cross validation (Appendix 4H) was also explored as a means to investigate the choice between models used in the variogram. The analysis indicated that kriging is relatively robust to the choice of the model. As well, the method presented statistical measures that could be used in determining the choice. During this discussion, it was pointed out that the Gaussian model should be avoided, with either the spherical or exponential model preferred. In particular, Gaussian models can tend to render the kriging solution unstable under certain conditions.

**Kriging Algorithm**

The chair reiterated the issue of how best to define the boundaries of the snow crab habitat for use in the kriging algorithm. As will be shown below, ordinary kriging is appropriate as long as changes to the survey boundaries are implemented. However, 'Kriging with External Drift (KED)' (Appendix 4I) which undertakes kriging using external variables to 'weight' the points used to define the response surface has considerable promise in making use of variables which define snow crab habitat in kriging. The utility of KED was illustrated by the use of bottom depth as a proxy for crab habitat (Appendix 4J)

There was discussion on the merits of point versus block kriging. It was pointed out that numerical procedures are required in the estimation of crab densities using block kriging, as is done for 4T. The size of the grid to use in the kriging is important; a 2 x 2 nm grid was considered adequate (Appendix 4K).

During this discussion, it was pointed out that it was important to check whether or not different geostatistical software packages produce the same result. In a comparison exercise conducted during the workshop, the results of different software programs used by participants were crosschecked and found to produce the same results. The 2003 commercial male survey data for Sector one was used with a Global variogram

analyses montrent que le krigeage est relativement robuste par rapport aux changements à l'intérieur de ces paramètres (appendice 4G).

On explore aussi la possibilité de recourir à la contre-validation pour le choix du modèle approprié pour le variogramme. L'analyse montre que le krigeage est relativement robuste par rapport au choix du modèle. De même, la méthode présente des mesures statistiques qu'on pourrait utiliser pour déterminer ce choix. Durant cette discussion, on fait remarquer qu'il faudrait éviter le modèle gaussien et accorder la préférence au modèle sphérique ou exponentiel. Les modèles gaussiens peuvent, en particulier, avoir tendance à rendre la solution en matière de krigeage instable dans certaines conditions.

**Algorithme du krigeage**

Le président ramène sur le tapis la question consistant à savoir comment le mieux définir les limites de l'habitat du crabe des neiges pour les utiliser dans l'algorithme du krigeage. Comme on le verra plus loin, le krigeage ordinaire (KO) est indiqué du moment qu'on apporte des changements aux limites du relevé. La méthode du KDE (appendice 4I) faisant appel à des variables externes pour « pondérer » les points utilisés afin de définir la surface de réponse est énormément prometteuse, parce qu'elle est fondée sur des variables qui définissent l'habitat du crabe des neiges dans le cadre du krigeage. L'utilisation de la profondeur au fond comme approximation de l'habitat du crabe illustre l'utilité du KDE (appendice 4J).

On discute des mérites du krigeage de points par opposition au krigeage de blocs. On fait remarquer qu'il faut utiliser des procédures numériques pour l'estimation des densités de crabes en faisant appel de krigeage de blocs, comme on l'a fait pour 4T. La taille de la grille à utiliser dans le cadre du krigeage est importante; on considère suffisante une grille de 2 x 2 NM (milles nautiques) (appendice 4K).

Durant cette discussion, on fait remarquer qu'il est important de vérifier si différents progiciels géostatistiques produisent le même résultat. Dans le cadre d'un exercice de comparaison effectué durant l'atelier, on contre-vérifie les résultats de différents progiciels utilisés par les participants et on constate qu'ils donnent les mêmes résultats. On utilise les données du relevé des mâles de taille commerciale de 2003 pour le secteur 1 et un

(spherical model) with the nugget (0.0075), range of influence (19.8892) and sill – nugget (2.2636) all defined. For three trials, the estimated biomass and estimated standard deviation (SD) were:

variogramme global (un modèle sphérique) et on définit la pépite (0,0075), la zone d'influence (19,8892) et le plateau – la pépite (2,2636). Pour trois essais, la biomasse estimée et l'écart-type (ET) estimé sont les suivants :

| Trial<br>Essai | Wade / Marcotte | Bez | Conan | Number of Stations<br>Nombre de stations |
|---|---|---|---|---|
| 1 | 4852.7 (341) | 4 820.15 (341,4) | 4801 (363) | 36 |
| 2 | 3869 (569) | - | 3863 (587) | 22 |
| 3 | 2816 (410) | - | 2809 (435) | 22 + 22<br>pseudo zeros<br>pseudozéros |

Data collected from an experimental survey designed by Marine Geomatics Inc. and provided by Dr. G. Conan as part of the common exercise are used in trial 1. In trial 2, all data used for the exercise are from the DFO Science sampling design and stations used as part of the regular 2003 trawl survey. The data set used in trial 3 is the same as in trial 2, but with 22 additional stations added by DFO Science outside the polygon and assigned zero values.

On utilise dans le cadre de l'essai n° 1 des données collectées à partir d'un relevé expérimental qu'a mis au point Marine Geomatics Inc. et fournies par M. G. Conan, D.Ph., dans le cadre de l'exercice commun. Dans le cas de l'essai n° 2, toutes les données utilisées pour l'exercice proviennent du plan d'échantillonnage et des stations employés par Sciences – MPO dans le cadre du relevé au chalut régulier de 2003. L'ensemble de données utilisé pour l'essai n° 3 est le même que celui employé pour l'essai n° 2, mais Sciences – MPO y ont ajouté 22 stations à l'extérieur du polygone et attribué des valeurs nulles.

During the review of these proceedings subsequent to the workshop, it was pointed out that the calculations reported by Wade / Marcotte did not appear to be based upon the DFO Science (Moncton) software. This situation was clarified. The DFO software, called MPOGEOS, consists of a MATLAB front end (user interface) that accesses two well established geostatistical libraries, GSLIB and COKRI. GSLIB is available as part of the book "GSLIB: Geostatistical Software Library and User's Guide" by Clayton Deutsch and Professor Andre Journel which is available from Oxford University Press. At the October workshop, as Marcotte was the author of COKRI, rather than conducting the test using the front end of MPOGEOS, he used COKRI directly.

Durant l'examen du présent compte rendu de l'atelier, on fait remarquer que les calculs présentés par Wade et Marcotte ne semblaient pas reposer sur le logiciel de Sciences – MPO (de Moncton). La situation a été éclaircie. Le logiciel du MPO, baptisé MPOGEOS, se compose d'un frontal MATLAB (une interface-utilisateur) qui donne accès à deux bibliothèques géostatistiques bien établies, GSLIB et COKRI. GSLIB est disponible à l'intérieur de l'ouvrage « GSLIB: Geostatistical Software Library and User's Guide », de Clayton Deutsch et du professeur Andre Journel, qu'on peut obtenir d'Oxford University Press. À l'atelier d'octobre, étant donné qu'il était l'auteur de COKRI, plutôt que d'effectuer l'essai à l'aide du frontal MPOGEOS, Marcotte a utilisé directement COKRI.

The test was rerun by Wade using MPOGEOS. His calculations are provided below and include a comparison to the results of the workshop. Marcottee independently conducted the same test and obtained the same results as Wade. The differences are very slight, which confirms the

Wade refait l'essai à l'aide du MPOGEOS. Ses calculs, qui figurent ci-dessous, incluent une comparaison avec les résultats de l'atelier. Marcotte effectue indépendamment le même essai et obtient les mêmes résultats que Wade. Les différences sont très minimes, ce qui

conclusion of the workshop that the software tested are producing essentially the same estimates.

confirme la conclusion des participants à l'atelier voulant que le logiciel mis à l'essai produit à l'heure actuelle essentiellement les mêmes estimations.

| Trial | Wade | Marcotte | Bez | Conan | Number of Stations |
|---|---|---|---|---|---|
| Essai | | | | | Nombre de stations |
| 1 | 4801 (341) | 4852,7 (341) | 4820,15 (341,4) | 4801 (363) | 36 |
| 2 | 3859 (571) | 3869 (569) | - | 3863 (587) | 22 |
| 3 | 2789 (404) | 2816 (410) | - | 2809 (435) | 22 + 22 pseudo zeros pseudozéros |

**Pseudo Zeros**

The chair reiterated that pseudo zeros have been used to delimit the boundaries of the survey but the question to the group is whether or not this is an appropriate means to define these boundaries. If not, what other means are available? Can environmental factors related to snow crab preferred habitat such as temperature and / or depth be used and if so, how?

**Pseudozéros**

Le président rappelle qu'on a utilisé des pseudozéros pour délimiter les limites du relevé; la question adressée au groupe consiste cependant à savoir s'il s'agit d'un moyen approprié de définir ces limites. Si ce n'est pas le cas, dispose-t-on d'autres moyens de le faire? Peut-on utiliser des facteurs environnementaux reliés à l'habitat préféré du crabe des neiges comme la température et/ou la profondeur et, si oui, comment les employer?

**Geographically Delimited Estimates of Biomass**

The chairman reiterated that the issue here was the estimation of biomass and its associated variance at the local scale. He queried if the entire survey area were divided into smaller local areas and the biomass of each evaluated, how would the resulting summed estimates compare to the globally determined biomass?

**Estimations délimitées géographiquement de la biomasse**

Le président rappelle que la question en jeu ici est l'estimation de la biomasse et de la variance qui y est associée à l'échelle locale. Il demande, si toute la zone du relevé est divisée en zones locales plus petites et si la biomasse de chacune est évaluée, comment la somme des estimations qui en découlent se compare-t-elle à la biomasse déterminée globalement?

## RESOLUTION OF ISSUES

**Impact of Spatial and Temporal Heterogeneity in Sampling Distribution**

To iterate the issue as outlined in the remit,

*The number of sampling sites has increased over the survey time series (nearly doubled since 1988). As well, the*

## RÈGLEMENT DES POINTS

**Répercussions de l'hétérogénéité spatiale et temporelle sur la distribution d'échantillonnage**

Pour reprendre les points dont le mandat fournit un aperçu,

*Le nombre de lieux d'échantillonnage a augmenté au cours de la série chronologique de relevés (il a presque*

*relative spatial distribution of this sampling effort has changed over time. What are appropriate statistical approaches to be used to ensure comparability in biomass estimates across the times series? Would this have consequences for the allocation of future sampling effort?*

*doublé depuis 1988). La distribution spatiale relative de cet échantillonnage a également évolué au fil du temps. Quelles sont les méthodes statistiques appropriées à utiliser pour assurer la comparabilité entre les estimations de la biomasse d'une série chronologique à une autre? Cela aurait-il des conséquences sur la détermination des futurs efforts d'échantillonnage?*

The distribution of sampling effort in the Southern Gulf of St. Lawrence has been motivated by competing objectives. The first is to ensure that the variograms are based upon sampling from all areas of the snow crab distributional field to avoid excessive weight being given to any one area in their construction. For this reason, when the variograms are being constructed, only one sample (first in chronological sample time) is chosen per survey sample cell. The second objective is to provide improved estimates of abundance, biomass and their distribution in certain areas, according to management demands. This latter objective has resulted in some areas being sampled more intensively compared to other areas. The first objective will be addressed here whereas the second will be addressed under the second issue.

La répartition de l'échantillonnage dans le sud du golfe du Saint-Laurent est motivée par des objectifs contradictoires. Le premier consiste à veiller à ce que les variogrammes soient fondés sur un échantillonnage de toutes les zones du champ de distribution du crabe des neiges pour éviter d'accorder trop de poids à l'une des zones dans le cadre de leur construction. Pour cette raison, lorsqu'on construit les variogrammes, on ne choisit qu'un seul échantillon (le premier dans le temps d'échantillonnage chronologique) par cellule d'échantillonnage du relevé. Le second objectif consiste à fournir de meilleures estimations de l'abondance, de la biomasse et de leur distribution dans certaines zones, suivant les exigences en matière de gestion. Ce dernier objectif entraîne un échantillonnage plus intensif de certaines zones comparativement à d'autres zones. On abordera ici le premier objectif, tandis qu'on traitera du second sous le point suivant.

The investigations of how the variograms of commercial male biomass changed over year and area indicated that while they appeared to be different, in fact, each variogram was more likely to be one realisation of a common underlying variogram for the Southern Gulf of St. Lawrence which may be changing slowly over time. When scaled for variance, the sill and range were positive linear functions of abundance while there was no obvious relationship between the nugget and slope and abundance. The nugget has components of both process and observation error, the latter of which may be possible to reduce through enhanced survey gear mensuration. Regarding the stability of the variogram over time, there did appear to be non-stationarity but there was also benefit in using information from previous years to enhance estimation of the current year's variogram. Explorations using annual smoothing averages of 2, 3, 5 and 10 years indicated that smoothing should not exceed a period of five years. While no criterion could be articulated to choose a smoothing period, three years seemed appropriate given the observed temporal trends and the

Les études sur la nature de l'évolution au fil des ans et des zones des variogrammes de la biomasse des mâles de taille commerciale indiquaient que même si les variogrammes semblaient différents, en fait, chacun d'eux était probablement davantage une réalisation d'un variogramme sous-jacent commun pour le sud du golfe Saint-Laurent qui peut changer lentement au fil du temps. Lorsqu'on les mettait à l'échelle pour tenir compte de la variance, on constatait que le plateau et la zone d'influence étaient des fonctions linéaires positives de l'abondance, tandis qu'il n'y avait aucune relation évidente entre la pépite et la pente et l'abondance. La pépite a des composantes à la fois d'erreur de traitement et d'erreur d'observation, cette dernière erreur pouvant être réduite grâce à un accroissement des dimensions de l'engin utilisé pour le relevé. En ce qui regarde la stabilité du variogramme au fil du temps, il ne semblait pas y avoir de stationnarité, mais on avait aussi avantage à utiliser des données d'années précédentes pour améliorer l'estimation du variogramme de l'année en cours. Les analyses faisant appel à des moyennes annuelles de

generation time of snow crab.

lissage de 2, 3, 5 et 10 ans indiquaient que ce dernier ne devrait pas dépasser une période de cinq ans. Même si l'on ne pouvait pas clairement définir de critère pour choisir une période de lissage, une période de trois ans semblait appropriée, compte tenu des tendances temporelles observées et du temps de génération du crabe des neiges.

It was therefore concluded that one variogram should be used and based upon the entire survey field of 4T snow crab (i.e.: snow crab fishing area 12, E, F and 19). This would be scaled by variance before being averaged over a three year period. Either a Spherical or Exponential model could be used but use of a Gaussian model was advised against, given stability problems experienced with the latter. Regarding specific choice of the model, cross validation as discussed at the workshop is a useful tool to assist in model selection as is sensitivity analysis. Overall, however, choice of the model was felt not to be a major issue as long as there is some investigation of goodness of fit and sensitivity.

On parvient donc à la conclusion qu'il faudrait utiliser un seul variogramme reposant sur toute la zone de relevé du crabe des neiges de 4T (c'est-à-dire les zones de pêche du crabe des neiges 12, E, F et 19), ce qu'on mettra à l'échelle suivant la variance avant d'en calculer la moyenne sur une période de trois ans. On pourrait utiliser un modèle sphérique ou exponentiel, mais on déconseille d'employer un modèle gaussien, étant donné les problèmes de stabilité éprouvés avec ce dernier. En ce qui a trait au choix précis du modèle, la contre-validation comme on en a discuté à l'atelier est un outil utile pour faciliter la sélection du modèle, comme l'est l'analyse de sensibilité. On estime cependant globalement que le choix du modèle n'est pas un point majeur, à condition d'étudier dans une certaine mesure la qualité de l'ajustement et la sensibilité.

Regarding the subsampling of one sample per cell, given that the variogram would be based upon the entire field, rescaled by variance, this might not now be necessary. However, it was recommended to calculate two variograms, one using all the data and another using the subsampled data to determine whether or not any differences are apparent.

En ce qui concerne le sous-échantillonnage d'un seul échantillon par cellule, étant donné que le variogramme serait fondé sur tout le champ, remis à l'échelle selon la variance, cela pourrait maintenant ne pas s'avérer nécessaire. On recommande cependant de calculer deux variogrammes, un à l'aide de toutes les données et un autre au moyen des données sous-échantillonnées, pour déterminer s'il y a des différences apparentes.

***Research Recommendation***

An analytical approach on the inclusion of habitat related information e.g. temperature and depth into the variogram using GLM was discussed. It was considered that these offer promise in improving descriptions of how variability changes with abundance and habitat.

***Recommandation en matière de recherche***

On discute d'une approche analytique de l'inclusion de données reliées à l'habitat comme la température et la profondeur à l'intérieur du variogramme à l'aide d'un MLG. On considère que ces données sont prometteuses pour l'amélioration des descriptions de la nature de l'évolution de la variabilité suivant l'abondance et l'habitat.

**Kriging Algorithm and Spatial Boundaries of Snow Crab Habitat**

To iterate the issue as outlined in the remit,

**Algorithme du krigeage et limites spatiales de l'habitat du crabe des neiges**

Pour reprendre les questions dont le mandat fournit un aperçu,

*Could information on snow crab habitat (e.g. water temperature, water depth, or combinations of these) be used to guide assumptions on spatial boundaries and if so, what would be the most appropriate analytical approach?*

*Pourrait-on utiliser des données sur l'habitat du crabe des neiges (comme la température de l'eau, la profondeur de cette dernière ou des combinaisons de ces données) pour orienter des hypothèses sur les limites spatiales et, si oui, quelle serait l'approche analytique la plus appropriée?*

For management issues to be addressed at the Gulf scale, global, ordinary kriging as implemented in 2004 was considered appropriate. Kriging with External Drift (KED) should be used and compared with ordinary kriging to see how external information could be used but this would be on an exploratory basis to learn how to use the approach and better understand its properties.

Pour les questions de gestion à aborder à l'échelle du golfe, on considère approprié le KO tel qu'il a été appliqué en 2004. On devrait utiliser le KDE et le comparer au KO pour voir comment il serait possible d'employer des données externes, mais cela se ferait de façon exploratoire pour apprendre de quelle(s) façon(s) se servir de l'approche et mieux comprendre ses propriétés.

For management issues of a local nature, KED was felt to be preferred as long as the appropriate search area was used. It was considered that KED might be more appropriate at smaller scales as the underlying crab – habitat associations in these restricted areas would be better described by this approach.

Pour les questions de gestion à caractère local, on estime préférable le KDE, à condition d'utiliser la zone de recherche appropriée. On considère que le KDE pourrait être davantage indiqué à des échelles plus petites, étant donné que cette approche décrirait mieux les associations sous-jacentes crabe – habitat à l'intérieur de ces zones restreintes.

Regarding the survey boundaries, the inside boundary should be set at 36.5 m while the outside boundary should be set at 200m. Expansion of the survey boundaries will necessitate distribution of current sampling effort into the inshore. New sampling cells will be required and sampling allocated to these as per the available budget. There are a number of ways to do this that will require further exploration.

En ce qui concerne les limites du relevé, la limite intérieure devrait être fixée à 36,5 m, tandis qu'on devrait en fixer la limite extérieure à 200 m. L'élargissement des limites du relevé nécessitera une répartition de l'effort d'échantillonnage actuel à l'intérieur des eaux côtières. Il faudra de nouvelles cellules d'échantillonnage et l'effort d'échantillonnage requis pour ces cellules dépendra du budget disponible. On peut utiliser un certain nombre de méthodes à cette fin, qui exigeront une étude plus poussée.

Having a larger survey area will also require careful treatment of the data to allow comparison with the historical time series. In the short term, it will be necessary to compare estimates based upon the 'old' survey area in the future surveys with those from the historical time series. As well, it will be useful to monitor the ratio of the abundance in the old and expanded areas in future surveys to see how stable this ratio is. If it is stable, then this information could be used to better interpret past surveys. If it is not, examination of the underlying biological processes will again lead to more insight in the historical population trends.

Le fait d'avoir une zone de relevé plus large exigera aussi un traitement minutieux des données pour permettre des comparaisons avec les séries chronologiques passées. À court terme, il faudra comparer des estimations fondées sur l'« ancienne » zone du relevé à l'occasion des futurs relevés à celles établies à partir des séries chronologiques historiques. De même, il sera utile de surveiller le ratio de l'abondance à l'intérieur de l'ancienne zone et de la zone élargie au moment des futurs relevés pour voir dans quelle mesure ce ratio est stable. S'il est stable, on pourrait alors utiliser cette information pour mieux interpréter les relevés passés. S'il est instable, un examen des processus biologiques sous-jacents permettra encore une fois de se faire une meilleure idée des tendances démographiques historiques.

***Research Recommendations***

As noted above, global Kriging with External Drift (KED) with the appropriate search area due to the quality of the temperature data should be explored. As well, co-kriging shows promise in the inclusion of environmental information into the analysis. Finally, further development of a snow crab habitat index was encouraged as a possible secondary variable to be included in the kriging analysis.

**Pseudo Zeros**

To iterate the issue as outlined in the remit,

> *What are the implications of using pseudo zeros to delineate snow crab distribution on abundance and biomass estimates? How many pseudo zeros should be used and in which situations?*

Pseudo zeros were used to compensate for edge effects in the analysis, especially along the shallow depth areas of the survey area. It was agreed that the use of pseudo-zeros is not necessary given that the issue of edge effects will be addressed by 1) using an expanded survey design that covers more of the boundaries and 2) the Kriging with external drift (KED) technique accounts for shallow depths in the analysis as discussed above.

**Procedures to Produce Geographically Delimited Estimate of Biomass**

To iterate the issue as outlined in the remit,

> *Can methodologies be developed to address fisheries management issues which require knowledge of biomass and abundance in small areas within the total population distribution?*

As per the conclusions on the first issue above, a global variogram should be used for biomass estimation in the Gulf of St. Lawrence. Local kriging can still be used for smaller areas to

***Recommandations en matière de recherche***

Comme il est souligné ci-dessus, on devrait analyser le KDE global avec la zone de recherche appropriée en raison de la qualité des données sur la température. De même le co-krigeage paraît prometteur sur le plan de l'inclusion de données environnementales dans l'analyse. Enfin, on est en faveur de l'élaboration plus poussée d'un indice de l'habitat du crabe des neiges comme variable secondaire possible à inclure dans l'analyse du krigeage.

**Pseudozéros**

Pour reprendre les points dont le mandat fournit un aperçu,

> *Quelles sont les conséquences de l'utilisation de pseudozéros pour délimiter la distribution du crabe des neiges sur les estimations de l'abondance et de la biomasse? Combien de pseudozéros devrait-on utiliser et dans quelles situations?*

On a utilisé des pseudozéros pour compenser les effets de lisière dans l'analyse, surtout le long des zones de faible profondeur de la zone du relevé. On convient que l'utilisation de pseudozéros n'est pas nécessaire, étant donné que 1) l'on a abordé la question des effets de lisière en utilisant un plan de relevé élargi qui englobe une plus grande partie des limites et que 2) la technique du KDE rend compte des petits fonds dans l'analyse, comme il en a été question ci-dessus.

**Procédures à suivre pour produire une estimation géographiquement délimitée de la biomasse**

Pour reprendre la question dont le mandat fournit un aperçu,

> *Peut-on mettre au point des méthodes afin de traiter de questions en matière de gestion des pêches qui exigent une connaissance de la biomasse et de l'abondance dans de petites zones à l'intérieur de la zone de distribution totale de la population?*

Suivant les conclusions tirées au sujet du point ci-dessus, on devrait utiliser un variogramme global pour l'estimation de la biomasse dans le golfe du Saint-Laurent. On peut encore employer

address specific management questions. It is expected that the estimates based upon the local kriging will exhibit higher variance which is to be expected and should be reported. Management activities, to be precautionary, should then take this uncertainty into account.

***Research Recommendation***

It would be useful to undertake conditional simulation to map the probability of changes in the resource. Specifically, probability maps of the incidence of soft shell crab or showing the probability of crab densities greater than a given level as predicted by the previous year's survey would be useful.

**Additional Issue on Sex Ratios in 2004 Survey**

Additional to the remit, a question was asked on the treatment of very high female / male abundance ratios in the survey, as observed in 2005. It was pointed out that this was likely due to sampling variation and the decision to subtract a portion of the landings from the commercial males. It was not an issue to be resolved on the agenda of this meeting.

**CONCLUDING REMARKS**

The chair thanked the participants for a productive meeting. The discussion covered a wide range of topics pertinent to the issues raised in the meeting remit and provided a number of recommendations on resolution of these issues as well as further research in these areas.

le krigeage local pour de plus petites zones afin de répondre à des questions précises en matière de gestion. On prévoit que les estimations fondées sur le krigeage local indiqueront une variance plus élevée à laquelle il faut s'attendre et qu'on devrait signaler. Il faudrait ensuite tenir compte de cette incertitude dans le cadre des activités de gestion, si l'on veut qu'elles soient axées sur des mesures de précaution.

***Recommandation en matière de recherche***

Il serait utile d'effectuer une simulation conditionnelle pour cartographier la probabilité de changements à l'intérieur de la ressource. Il serait tout spécialement utile de disposer de cartes des probabilités de la fréquence du crabe à carapace molle ou indiquant la probabilité de densités de crabes supérieures à un niveau donné prévu par le relevé de l'année précédente.

**Point additionnel sur les rapports des sexes dans le relevé de 2004**

En plus des questions mentionnées qui relèvent du mandat, on en pose une sur le traitement des ratios d'abondance femelles/mâles très élevés dans le relevé, comme on l'a observé en 2005. On fait remarquer que cela était probablement dû à la variation de l'échantillonnage et à la décision de soustraire une portion des débarquements des mâles de taille commerciale. Ce n'est pas un point à régler à l'ordre du jour de la présente réunion.

**MOT DE LA FIN**

Le président remercie les participants d'avoir fait de l'atelier une réunion productive. La discussion a englobé une vaste gamme de sujets pertinents liés au mandat des participants à la réunion et a fourni un certain nombre de recommandations au sujet du règlement de ces points et à propos de recherches plus poussées dans ces domaines.

## APPENDICES

### Appendix 1. List of Participants / Appendice 1. Liste des participants

| Surname / Prénom | Name/Nom | Institution / Organization | E-Mail/Courriel |
|---|---|---|---|
| Monique | Baker | DFO, GFC, Moncton, NB | Bakerm@dfo-mpo.gc.ca |
| Pierre | Bélanger | DFO, GFC, Moncton, NB | Belangerp@dfo-mpo.gc.ca |
| Nicolas | Bez | iRD, Sète, France | Nicola.bez@ifremer.fr |
| Michel | Biron | DFO, GFC, Moncton, NB | Bironm@dfo-mpo.gc.ca |
| Hugo | Bourdages | DFO, IML, Mont-Joli, Qc | Bourdagesh@dfo-mpo.gc.ca |
| Jae | Choi | DFO, BIO, Dartmouth, NS | Choij@mar.dfo-mpo.gc.ca |
| Gérard | Conan | Marine Geomatics Inc. | marinegeomatics@primus.ca |
| Réjean | Dufour | DFO, IML, Mont-Joli, Qc | Dufourr@dfo-mpo.gc.ca |
| Marc | Lanteigne | DFO, GFC, Moncton, NB | Lanteignem@dfo-mpo.gc.ca |
| Julie | Marcil | DFO, GFC, Moncton, NB | Marcilj@dfo-mpo.gc.ca |
| Denis | Marcotte | École Polytechnique de Montréal | Denis.marcotte@polymtl.ca |
| Monique Anne | Morin | NBDFA, Fredericton, NB | Monique.anne.morin@gnb.ca |
| Mikio | Moriyasu | DFO, GFC, Moncton, NB | Moriyasu@dfo-mpo.gc.ca |
| Robert - *Chair* | O'Boyle | DFO, BIO, Dartmouth, NS | Oboyler@mar.dfo-mpo.gc.ca |
| Bruce | Osborne | NSDFA, Halifax, NS | Osbornbd@gov.ns.ca |
| Stephen | Smith | DFO, BIO, Dartmouth, NS | Smithsj@mar.dfo-mpo.gc.ca |
| Tobie | Surette | DFO, GFC, Moncton, NB | Surettet@dfo-mpo.gc.ca |
| Elmer | Wade | DFO, GFC, Moncton, NB | Wadee@dfo-mpo.gc.ca |

## Appendix 2. Remit

### Background

Snow crab in the southern Gulf of St. Lawrence has been commercially harvested since the mid-1960s. In support of stock harvesting advice, stock assessments have been conducted each spring prior to the fishery. Since 1988, the assessment has depended upon fishery independent research surveys which are partially funded by the fishing industry. The survey design (e.g. area covered, sampling density, and survey vessel) has evolved over time to provide more accurate fishable biomass and recruitment abundance indices, which have been generated using a geostatistical (i.e. kriging) approach.

In 2004, an alternate analysis by a non-DFO expert using the same data but with different statistical assumptions highlighted issues that needed to be addressed, for example the method used to define the spatial boundary of the distribution of commercial-sized male snow crab and the surface assumed for biomass and abundance estimation. These assumptions have significant influence on the total biomass and recruitment abundance estimates used for the management of the fishery.

The workshop is to review current treatment of the survey data and explore methodologies that might lead to improved biomass estimates for management purposes.

### Objectives

The following objectives will be used to guide workshop activities:

1. Explore the impact of heterogeneity in the spatial and temporal distribution of sampling sites on the biomass estimates and variograms.

The number of sampling sites has increased over the survey time series (nearly doubled since 1988). As well, the relative spatial distribution of this sampling effort has changed over time. What are appropriate statistical approaches to be used

## Appendice 2. Mandat

### Contexte

Le crabe des neiges dans le sud du golfe Saint-Laurent est commercialement exploité depuis le milieu des années 60. Pour appuyer les avis sur l'exploitation du stock, on effectue des évaluations de ce dernier chaque printemps avant la pêche. Depuis 1988, l'évaluation dépend de relevés de recherche effectués indépendamment de l'industrie de la pêche et qui sont en partie financés par ce secteur d'activité. On a modifié le plan de relevé (la zone couverte, la densité d'échantillonnage, le bateau affecté au relevé, etc.) au fil du temps de façon à fournir des indices plus fiables de la biomasse exploitable et de l'abondance des recrues établis à l'aide d'une méthode géostatistique (c'est-à-dire le krigeage).

En 2004, une autre analyse effectuée par un expert étranger au MPO ayant utilisé les mêmes données, mais suivant des hypothèses statistiques différentes a mis en lumière des points qu'il faut aborder, notamment la méthode employée pour définir la limite spatiale de la distribution du crabe des neiges mâle de taille commerciale et la surface présumée pour l'estimation de la biomasse et de l'abondance. Ces hypothèses ont une influence significative sur les estimations de la biomasse totale et de l'abondance des recrues qu'on utilise pour la gestion de la pêche.

Les participants à l'atelier doivent examiner le traitement actuel des données de relevé et étudier à fond des méthodes qui pourraient mener à une amélioration des estimations de la biomasse à des fins de gestion.

### Objectifs

On utilisera les objectifs suivants pour orienter les activités des participants à l'atelier :

1. Analyser les répercussions de l'hétérogénéité à l'intérieur de la distribution spatiale et temporelle des lieux d'échantillonnage sur les estimations et les variogrammes de la biomasse.

Le nombre de lieux d'échantillonnage a augmenté au fil des séries chronologiques de relevé (il a presque doublé depuis 1988). De même, la distribution spatiale relative de cet échantillonnage a évolué au fil du temps. Quelles

to ensure comparability in biomass estimates across the times series? Would this have consequences for the allocation of future sampling effort?

2. Investigate the kriging algorithm, particularly in regards to establishing spatial boundaries of snow crab habitat.

Could information on snow crab habitat (e.g. water temperature, water depth, or combinations of these) be used to guide assumptions on spatial boundaries and if so, what would be the most appropriate analytical approach?

3. In cases where pseudo zeros are used, considerations are made on their number and location.

What are the implications of using pseudo zeros to delineate snow crab distribution on abundance and biomass estimates? How many pseudo zeros should be used and in which situations?

4. Investigate procedures using a subset of the survey data to produce geographically delimited estimate of biomass for management purposes

Can methodologies be developed to address fisheries management issues which require knowledge of biomass and abundance in small areas within the total population distribution?

**Products**

- CSAS Proceedings summarizing workshop discussion
- CSAS Research Document providing the technical details of the treatment of the survey information

**Participation**

- DFO Science
- DFO Fisheries Management
- Provinces NS & NB

sont les méthodes statistiques appropriées à utiliser pour assurer la comparabilité des estimations de la biomasse d'une série chronologique à une autre? Cela aurait-il des conséquences sur l'allocation des futurs efforts d'échantillonnage?

2. Étudier l'algorithme du krigeage, en particulier par rapport à l'établissement de limites spatiales de l'habitat du crabe des neiges.

Pourrait-on utiliser les données sur l'habitat du crabe des neiges (comme la température de l'eau, la profondeur de cette dernière ou des combinaisons de ces données) pour orienter les hypothèses au sujet des limites spatiales et, si oui, quelle serait la méthode analytique la plus appropriée?

3. Dans les cas où on utilise des pseudozéros, examiner leur nombre et leurs emplacements.

Quelles sont les conséquences de l'utilisation de pseudozéros visant à délimiter la distribution du crabe des neiges sur les estimations de l'abondance et de la biomasse? Combien de pseudozéros devrait-on utiliser et dans quelles situations?

4. Étudier des procédures faisant appel à un sous-ensemble des données de relevé pour produire une estimation géographiquement délimitée de la biomasse à des fins de gestion.

Peut-on mettre au point des méthodes pour répondre à des questions en matière de gestion des pêches qui exigent une connaissance de la biomasse et de l'abondance dans de petites zones à l'intérieur de la distribution totale de la population?

**Produits**

- Compte rendu du SCCS résumant les discussions des participants à l'atelier
- Document de recherche du SCCS fournissant les détails techniques du traitement des données de relevé

**Participants**

- MPO – Sciences
- MPO – Gestion des pêches
- Provinces de la Nouvelle-Écosse et du

- External Reviewers

Nouveau-Brunswick

- Examinateurs externes

**Appendix 3. Agenda**

**10 October – Monday**

Travel

**11 October – Tuesday**

09:00 - 09:30 Welcome and Introduction (Chair)

09:30 - 10:30 Working Paper on Workshop Issues (E. Wade)
10:30 - 10:45 Break
10:45 - 12:00 Discussion & Reviewer Comments

12:00 – 13:00 Lunch

13:00 - 14:30 Working Paper on Workshop Issues (G. Conan)
14:30 - 15:30 Discussion
15:30 - 15:45 Break
15:45 - 16:30 Presentation on Survey Treatment of 4VW Snow Crab (J. Choi)

16:30 - 17:15 Discussion

**12 October - Wednesday**

09:00 – 09:30 Synopsis of First Day's Discussion and Issues
09:30 - 11:00 Discussion & Workshop Planning for Issues
- Impact of Heterogeneity of Temporal / Spatial Distribution
- Kriging Algorithm
- Pseudo – Zeros
- Procedure to use Survey Subset

11:00 – 12:00 Exploratory Analyses by Participants with Presentations as needed

12:00 – 13:00 Lunch

13:00 – 17:00 Exploratory Analyses by Participants with Presentations as needed

**13 October - Thursday**

09:00 – 12:00 Review of Analyses & Discussion & Resolution on Issues

12:00 – 13:00 Lunch

**Appendice 3. Ordre du jour**

**10 octobre – Lundi**

Déplacements

**11 octobre – Mardi**

De 9 h à 9 h 30 Mot de bienvenue et présentation des participants (le président)
De 9 h 30 à 10 h 30 Document de travail sur les points de l'atelier (E. Wade)
De 10 h 30 à 10 h 45 Pause
De 10 h 45 à 12 h Discussion et commentaires des examinateurs

De 12 h à 13 h Dîner

De 13 h à 14 h 30 Document de travail sur les points de l'atelier (G. Conan)
De 14 h 30 à 15 h 30 Discussion
De 15 h 30 à 15 h 45 Pause
De 15 h 45 à 16 h 30 Exposé sur le traitement des données de relevé sur le crabe des neiges de 4VW (J. Choi)
De 16 h 30 à 17 h 15 Discussion

**12 octobre – Mercredi**

De 9 h à 9 h 30 Synthèse de la première journée et des points discutés
De 9 h 30 à 11 h Planification de la discussion et de l'atelier sur les points à l'étude
- Répercussions de l'hétérogénéité à l'intérieur de la distribution temporelle et spatiale
- Algorithme du krigeage
- Pseudozéros
- Procédure à suivre pour utiliser un sous-ensemble de données de relevé

De 11 h à 12 h Analyses exploratoires par les participants et exposés au besoin

De 12 h à 13 h Dîner

De 13 h à 17 h Analyses exploratoires par les participants et exposés au besoin

**13 octobre – Jeudi**

De 9 h à 12 h Examen des analyses et discussion et règlement des points

De 12 h à 13 h Dîner

| | | | |
|---|---|---|---|
| 13:00 – 17:00 | Synopsis of Meeting's Findings | De 13 h à 17 h de la réunion | Synthèse des conclusions |

**14 October - Friday (if Needed)**

**14 octobre – Vendredi (au besoin)**

| | | | |
|---|---|---|---|
| 09:00 – 12:00 | Synopsis of Meeting's Findings | De 9 h à 12 h de la réunion | Synthèse des conclusions |
| 12:00 – 13:00 | Adjournment | De 12 h à 13 h | Clôture |

## Appendix 4. Exploratory Analyses undertaken at the Meeting

### Appendix 4A. J. Choi. Use of Past Survey Information

Empirical variograms and the subsequent variogram models were potentially unstable, especially in cases where spatial distributions of snow crab were patchy or when analyses were conducted in a small area with limited spatial sample densities. A means of stabilizing these variograms was examined by using temporal moving averaging of varying window sizes. Using the 4T survey data, empirical variograms, normalized to unit variance, were fit to a spherical model. The parameters of the fit, range of influence, nugget and partial sill, were determined for a number of permutations of window size (1 to 10 years) and decay (uniform influence, linear decay with time (l), exponential decay with time (e)). Not all combinations of window size and decay type provided good solutions. In fact, many had singular solutions. Those solutions that were not singular are given in Figure 4A.

While no temporal drift in the parameters of the modeled annual variograms was expected, drift was observed in variogram range of influence, partial sill and nugget (the temporal drift of the nugget was however small relative to the range and partial sill). Inter-annual variability was also quite evident. To stabilize this temporal variability, a three-year running average of the empirical variograms, regardless of decay, was found to be optimal for the stabilization of the modeled variograms.

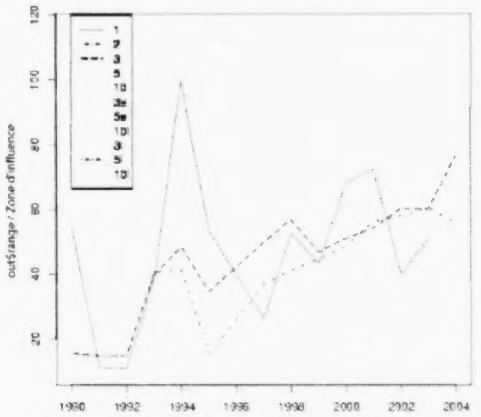
## Appendice 4. Analyses exploratoires entreprises à la réunion

### Appendice 4A. J. Choi. Utilisation de données de relevés antérieurs

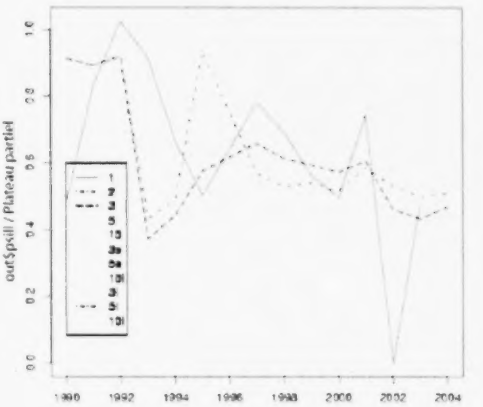
Les variogrammes empiriques et les modèles de variogrammes subséquents pouvaient être instables, surtout dans les cas où les distributions spatiales du crabe des neiges étaient inégales ou lorsque l'on effectuait des analyses dans une petite zone à l'intérieur de laquelle les densités d'échantillonnage dans l'espace étaient limitées. On a examiné un moyen de stabiliser ces variogrammes en utilisant des moyennes mobiles temporelles dont les tailles de fenêtre variaient. À l'aide des données du relevé de 4T, on a ajusté à un modèle sphérique les variogrammes empiriques, normalisés suivant la variance des unités. On a déterminé les paramètres de l'ajustement, la zone d'influence, la pépite et le tableau partiel, pour un certain nombre de permutations de taille de fenêtre (1 à 10 ans) et de décroissance (influence uniforme, décroissance linéaire avec le temps (l), décroissance exponentielle avec le temps (e)). Ce ne sont pas toutes les combinaisons de taille de fenêtre et de type de décroissance qui ont fourni de bonnes solutions. En fait, beaucoup présentaient des solutions singulières. Les solutions qui n'étaient pas singulières sont indiquées à l'intérieur de la figure 4A.

Même si on ne s'attendait à aucune dérive temporelle à l'intérieur des paramètres des variogrammes annuels modélisés, on a observé une dérive dans la zone d'influence, le plateau partiel et la pépite des variogrammes (la dérive temporelle de la pépite était cependant petite par rapport à la zone d'influence et au plateau partiel). La variabilité interannuelle était aussi assez évidente. Pour stabiliser cette variabilité temporelle, on a constaté qu'une moyenne mobile de trois ans des variogrammes empiriques, indépendamment de la décroissance, était la solution optimale pour la stabilisation des variogrammes modélisés.

a. Range of Influence / Zone d'influence




b. Sill /Plateau




c. Nugget / Pépite

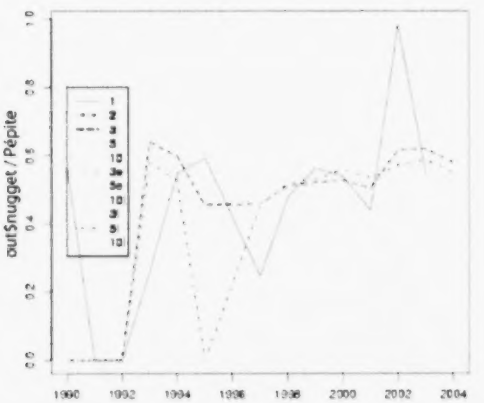



*Figure 4A Explorations of changes in (a) variogram range, (b) sill, and (c) and nugget using different window sizes over years and decay assumptions; see text for legend descriptions.*

*Figure 4A Analyses des changements dans a) la zone d'influence, b) le plateau et c) la pépite à l'aide de différentes tailles de fenêtre au fil des ans et hypothèses au sujet de la décroissance; voir le texte pour la description des légendes.*

**Appendix 4B. E. Wade. The Relationship Between Nugget, Range of Influence and Sill**

The parameters of the variograms of commercial population were compared with the commercial snow crab abundance for each year. These analyses were done using the global southern Gulf of St. Lawrence dataset. Results show that a significant correlation exists between the snow crab abundance and the range of influence (Fig.

**Appendice 4B. E. Wade. La relation entre la pépite, la zone d'influence et le plateau**

On a comparé pour chaque année les paramètres des variogrammes de la population de taille commerciale avec l'abondance du crabe des neiges de taille commerciale. Ces analyses ont été effectuées à l'aide de tout l'ensemble de données du sud du golfe du Saint-Laurent. Les résultats montrent qu'il existe une corrélation

4Ba). This suggests that as abundance increases, an expansion of the range of the concentration occurs.

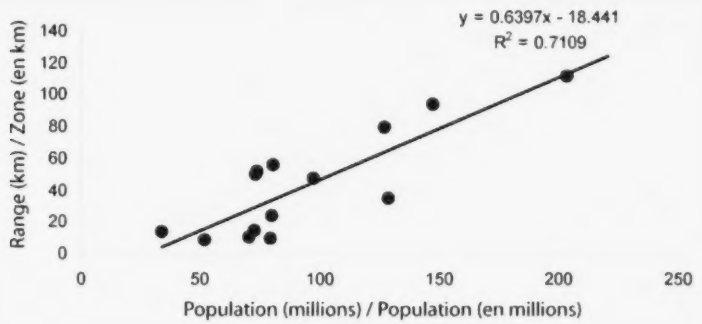



*Fig. 4Ba. Abundance of commercial males versus range of influence from variograms.*

significative entre l'abondance du crabe des neiges et la zone d'influence (fig. 4Ba), ce qui laisse entendre qu'il y a élargissement de la zone de concentration à mesure que l'abondance augmente.

*Fig. 4Ba. Effet de l'abondance des crabes mâles de taille commerciale sur la zone d'influence d'après les variogrammes*

A correlation also exists between the abundance of commercial crab and the sill parameter from the variogram analysis (Fig. 4Bb).

Il existe une corrélation entre l'abondance du crabe de taille commerciale et le paramètre plateau d'après l'analyse des variogrammes (fig. 4Bb).

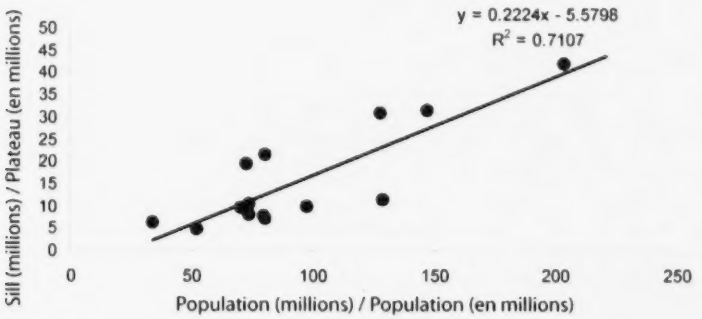



*Fig. 4Bb. Abundance of commercial males versus sill parameter from variograms.*

*Fig. 4Bb. Effet de l'abondance des crabes mâles de taille commerciale sur le paramètre plateau d'après les variogrammes*

And finally a positive correlation exists between the range of influence parameter and the sill parameter from the variograms (Fig. 4Bc).

Finalement, il existe en plus une corrélation positive entre le paramètre zone d'influence et le paramètre plateau d'après les variogrammes (fig. 4Bc).

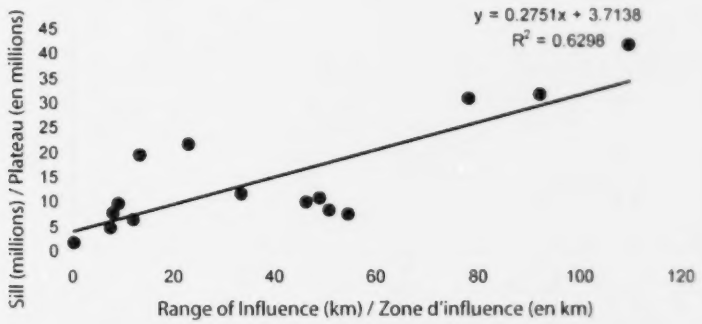



*Fig. 4Bc. Range of influence versus sill parameter from variograms.*

*Fig. 4Bc. Effet de la zone d'influence sur le paramètre plateau d'après les variogrammes*

**Appendix 4C. T. Surette. The Slope at the Origin of a variogram**

For a spherical variogram model given by

$$g(h)=\begin{cases} C_0+C_1\left(\dfrac{3h}{2a}-\dfrac{1}{2}\left(\dfrac{h}{a}\right)^3\right), & h\le a \\ C_0+C_1, & h>a \end{cases}$$

the slope at the origin $h=0$ is given by

$$g'(0)=\frac{3C_1}{2a}$$

There was no relationship observed between the slope at the origin of the variograms examined and the overall abundance estimate (Fig. 4C).

**Appendice 4C. T. Surette. La pente à l'origine d'un variogramme**

Pour un modèle de variogramme sphérique donné par

$$g(h)=\begin{cases} C_0+C_1\left(\dfrac{3h}{2a}-\dfrac{1}{2}\left(\dfrac{h}{a}\right)^3\right), & h\le a \\ C_0+C_1, & h>a \end{cases}$$

la pente à l'origine $h=0$ est fournie par

$$g'(0)=\frac{3C_1}{2a}$$

On n'observe pas de relation entre la pente à l'origine des variogrammes examinés et l'estimation globale de l'abondance (fig. 4C).

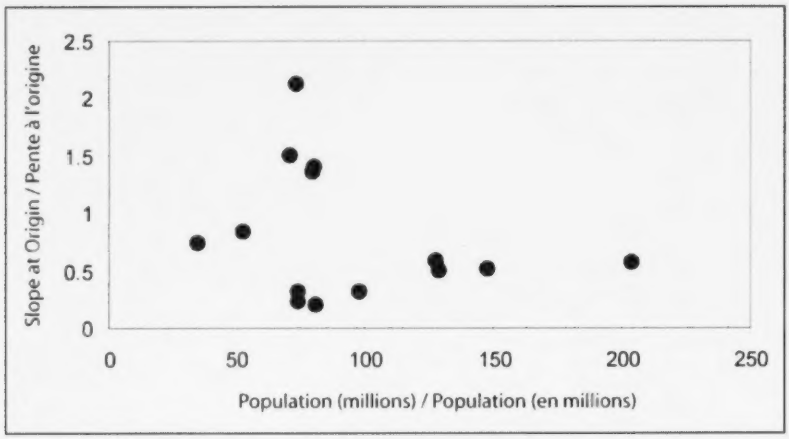



*Fig. 4C. Slope at the origin of variograms versus abundance estimates.*

*Fig. 4C. Pente à l'origine des variogrammes par opposition aux estimations de l'abondance*

**Appendix 4D. N. Bez. Year to Year Variability of the Variograms**

To assess the variability of the variograms' shape from one survey to the next (i.e. from one year to the next), we compared the variograms obtained for the commercial density (raw and log transformed) and for the bathymetry, a variable for which little (if no) year to year variability is expected (Figure 4D1). To allow comparing variograms regardless of the level of the variance, relative variograms were considered (sill =1).

Clearly, at the log-scale, commercial crab density has a very similar spatial structure from year to year. The annual variation, whose consequences in terms of kriging interpolation and estimation variance, will be negligible. In addition, the variation observed in the variograms of the log transformed data are of the same order of magnitude than those observed for bathymetry.

Even though the picture is less clear for crab density at the raw scale, this goes towards the use of a single variograms shape, or a slowly variable variogram shape (see below and Appendix 4A), scaled to the annual or local variance.

**Appendice 4D. N. Bez. Variabilité d'une année à une autre des variogrammes**

Pour évaluer la variabilité de la forme des variogrammes d'un relevé au suivant (c'est-à-dire d'une année à la suivante), nous avons comparé les variogrammes obtenus pour la densité de crabes de taille commerciale (brute et logarithmique transformée) et pour la bathymétrie, une variable pour laquelle on s'attend à peu de variabilité (sinon à aucune) d'une année à une autre (figure 4D1). Pour permettre de comparer des variogrammes indépendamment du niveau de la variance, on a examiné des variogrammes relatifs (plateau = 1).

À l'échelle logarithmique, manifestement, la densité de crabes de taille commerciale a une structure spatiale très similaire d'une année à une autre. Par conséquent, comme la variabilité annuelle du variogramme est faible, l'interpolation du krigeage et la variance de l'estimation seront négligeables. En outre, la variation observée à l'intérieur des variogrammes des données logarithmiques transformées est d'une ampleur similaire à celle observée pour la bathymétrie.

Même si le portrait est moins clair pour la densité de crabes à l'échelle brute, cela mène à l'utilisation d'une forme de variogramme unique ou d'une forme de variogramme qui varie lentement (voir ci-dessous et l'appendice 4A), mise à l'échelle suivant la variance annuelle ou locale.

To account for slight changes of spatial structure parameters through time, three and five year running means were also considered (Figure 4D2). A five year running mean allows reducing the variogram fluctuations while detecting semi decadal changes (if they exist). Three year running means could also represent an interesting compromise, but remain impacted by individual variogram variability.

Pour rendre compte des légers changements des paramètres de la structure spatiale au fil du temps, on a aussi examiné des moyennes mobiles de trois et de cinq ans (figure 4D2). Une moyenne mobile de cinq ans permet de réduire les fluctuations des variogrammes pendant qu'on détecte des changements semi-décennaux (le cas échéant). Les moyennes mobiles de trois ans pourraient aussi représenter un compromis intéressant, mais demeurent influencées par la variabilité de chaque variogramme.

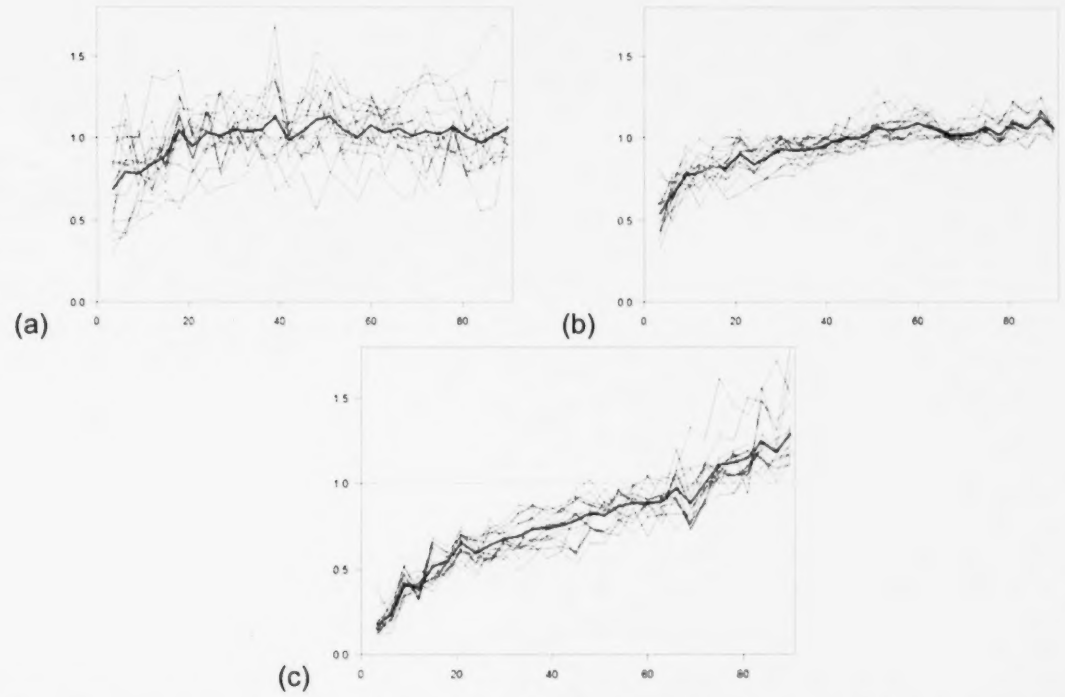
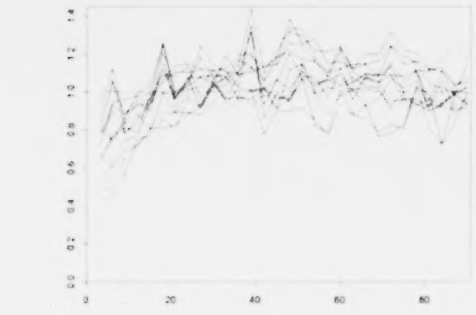
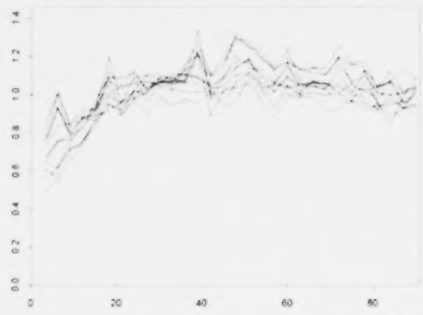



*Fig. 4D1 Annual standardized variograms (light continuous lines) and mean annual variogram (wide continuous line) for (a) commercial density, (b) log transformed commercial density, and (c) depth.*

*Fig. 4D1 Variogrammes normalisés annuels (fins traits continus) et variogramme annuel moyen (large trait plein) pour a) la densité de crabes de taille commerciale, b) la densité de crabes de taille commerciale logarithmique transformée et c) la profondeur.*



(a)



(b)

*Fig. 4D2 Running means of annual variograms. Raw commercial crab density. (a) 3 year running means (i.e. actual with the two previous years). (b) 5 year running means.*

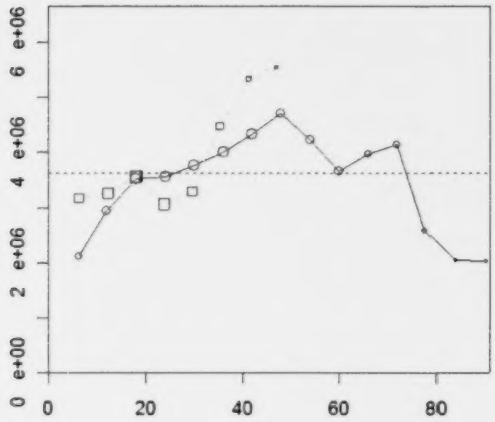
*Fig. 4D2 Moyennes mobiles de variogrammes annuels. Densité brute de crabes de taille commerciale. a) Moyennes mobiles de trois ans (c'est-à-dire données réelles plus les deux années précédentes). b) Moyennes mobiles de cinq ans.*

**Appendix 4E. N. Bez. Anisotropy of the Crab Distribution in the CB Subunit**

Figure 4E illustrates the investigated suspected anisotropy depicted by the variogram structure in the Cape Breton area. However, for the 2004 case at least, no clear difference exists between variograms computed along or across shelf direction and the concept of anisotropy is not strongly supported by the data.

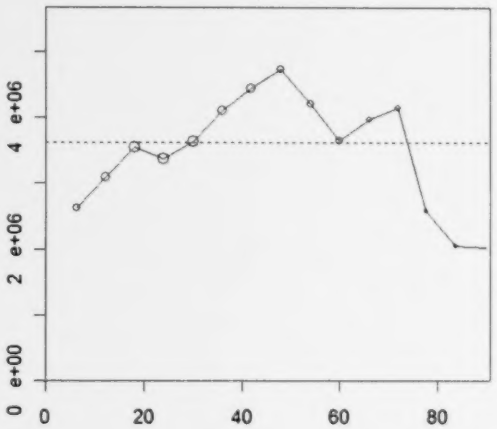
**Appendice 4E. N. Bez. Anisotropie de la distribution du crabe dans la sous-unité du C-B**

La figure 4E illustre l'anisotropie soupçonnée à l'étude que dépeint la structure des variogrammes dans la zone du Cap-Breton. Pour 2004 tout au moins, cependant, il n'existe aucune différence évidente entre les variogrammes calculés le long ou en travers de la direction de la plate-forme et les données n'étayent pas non plus solidement le concept d'anisotropie.




(a)

*Fig 4E. Evaluating anisotropy in CB subunit, 2004, commercial density.*
*(a) Variogram computed along (~40°) and across shelf (~130°). E –W direction = 0°.*
*(b) Omnidirectional variogram.*




(b)

*Fig. 4E. Évaluation de l'anisotropie dans la sous-unité du C-B, 2004, densité du crabe de taille commerciale*
*a) Variogramme calculé le long (~40°) et en travers de la plate-forme (~130°). Direction E –W = 0°.*
*b) Variogramme omnidirectionnel*

**Appendix 4F. S. Smith. Analysis of Spatial Data**

The standard geostatistical model can be written in the following linear model form which is more commonly used in statistical modelling (Cressie 1993, Bannerjee et al. 2004).

(1) $Y(s) = \mu(s) + \omega(s) + \varepsilon(s)$,

where Y (s) represents the measurement taken (e.g., number of crab) taken at spatial location s (i.e., latitude and longitude, northings and eastings, etc.). The mean structure is represented by μ(s) and can be a simple global mean or involve covariates, for example,

$$\mu(s) = x^T(s)\beta.$$

The ω(s) are realizations from a zero-mean stationary Gaussian spatial process which includes the partial sill ($\sigma^2$) and range ($\phi$) parameters. The ε(s) are uncorrelated pure error terms (nugget effect, $\tau^2$).

**Appendice 4F. S. Smith. Analyse des données spatiales**

On peut présenter par écrit le modèle géostatistique type sous la forme du modèle linéaire suivant qui est plus couramment utilisé en modélisation statistique (Cressie 1993, Bannerjee et coll., 2004).

(1) $Y(s) = \mu(s) + \omega(s) + \varepsilon(s)$,

où Y(s) représente la mesure prise (comme le nombre de crabes) à des points géographiques (c'est-à-dire des degrés de latitude et de longitude, des ordonnées et des abscisses, etc.). La structure moyenne est représentée par μ(s) et peut être une moyenne globale simple ou englober des covariables, par exemple,

$$\mu(s) = x^T(s)\beta.$$

Les ω sont des réalisations établies à partir d'un processus spatial gaussien stationnaire de moyenne zéro qui inclut les paramètres plateau partiel ($\sigma^2$) et zone d'influence ($\phi$). Les ε sont des termes d'erreur pure non corrélés (l'effet de pépite, $\tau^2$).

Given a sample of $Y(s_i)$, i = 1, . . . , n and $\mathbf{Y} = (Y(s_1, \ldots, Y(s_n))^T$, then we can write $\mathbf{Y}$ as a special case of the general linear model where,

(2) $\mathbf{Y}|\theta \sim N(\mathbf{X}\beta, \sigma^2\mathbf{H}(\phi) + \tau^2\mathbf{I})$.

The matrix **H** is a correlation matrix with elements $H_{ij} = \rho(s_i - s_j\ ;\ _)$ where $\rho$ is a valid isotropic correlation function indexed by parameter $\phi$. The $\theta$ term represents the vector of parameters to be estimated, i.e., $\beta$, $\sigma^2$, $\tau^2$ and $\phi$. Parameter estimates can be obtained using maximum likelihood (ML), restricted maximum likelihood (REML) or in a Bayesian analysis (Pinheiro and Bates 2000, Bannerjee et al. 2004). The linear (can also be non-linear) model structure will allow for standard statistical software available in SAS, SPLUS, R and WinBugs to be used to model relations between the mean structure and covariates. In addition, the mixed effects models in these packages will also allow for comparing different forms for variogram models as well as incorporating covariates into the variograms themselves.

***Application to Snow Crab Data***

Below, the SPLUS function lme() is used to illustrate how a standard statistical package can be used to analyze the spatial structure in snow crab survey data. Simple random effects models for the mean were fitted to commercial density data from surveys conducted in the 4T area for 2002, 2003 and 2004. Exponential and spherical models were compared using Akaike's Information Criterion (AIC). Observations from all areas in the 4T area were used to construct global variograms as per standard practice in the stock assessment although there was a considerable increase in variance (sill) from just using those observations from the Area 12 subunit.

The differences in AIC between the two models indicates that the exponential model fit the best although the improvement in fit was marginal in

À partir d'un un échantillon de $Y(s_i)$, i = 1, . . . , n et $\mathbf{Y} = (Y(s_1, \ldots, Y(s_n))^T$, nous pouvons ensuite représenter par écrit $\mathbf{Y}$ comme étant un cas spécial du modèle linéaire général où

(2) $\mathbf{Y}|\theta \sim N(\mathbf{X}\beta, \sigma^2\mathbf{H}(\phi) + \tau^2\mathbf{I})$.

La matrice **H** est une matrice de corrélation avec les éléments $H_{ij} = \rho(s_i - s_j\ ;\ _)$ dans laquelle $\rho$ est une fonction de corrélation isotropique valide indexée par le paramètre $\phi$. Le terme $\theta$ représente le vecteur de paramètre à estimer, c'est-à-dire $\beta$, $\sigma^2$, $\tau^2$ et $\phi$. On peut obtenir des estimations du paramètre à l'aide d'un maximum de vraisemblance (ML), au moyen d'un maximum de vraisemblance limité (REML) ou dans une analyse de Bayes (Pinheiro et Bates, 2000, Bannerjee et coll., 2004). La structure de modèle linéaire (qui peut aussi être non linéaire) permettra d'utiliser le logiciel statistique standard disponible dans SAS, SPLUS, R et WinBugs pour modéliser des relations entre la structure moyenne et les covariables. En plus, les modèles aux effets mixtes dans ces progiciels permettront aussi de comparer différentes formes pour des modèles de variogrammes et d'intégrer des covariables aux variogrammes eux-mêmes.

***Application aux données sur le crabe des neiges***

On utilise ci-dessous la fonction lme() de SPLUS pour illustrer comment l'on peut employer un logiciel statistique courant afin d'analyser la structure spatiale à l'intérieur de données de relevé du crabe des neiges. On a ajusté les modèles aux effets aléatoires simples établis pour la moyenne à des données sur la densité de crabes de taille commerciale tirées de relevés effectués dans la division 4T en 2002, en 2003 et en 2004. On a comparé des modèles exponentiel et sphérique à l'aide du critère d'information d'Akaike (AIC). On a utilisé des observations effectuées à partir de toutes les zones de la division 4T pour construire des variogrammes globaux suivant la pratique courante en évaluation des stocks, même s'il y a eu une augmentation énorme de la variance (du plateau) découlant du simple fait d'employer ces observations tirées de la sous-unité de la zone 12.

La différence à l'intérieur du critère d'information d'Akaike entre les deux modèles indique que c'est le modèle exponentiel qui offre le meilleur

2004 and 2002 (Table 4Fa).

Attention was drawn during the workshop to the similarity between the variogram for the commercial density and that for the depths observed in the survey. Mean gamma estimates for the same distance bin are compared in Figure 4Fa using data from 2004 where the strong linear pattern is evident. There is also appears to be a similar strong pattern for temperature (Fig. 4Fb).

The lme() software allows for including covariates such as depth into the variogram modeling. Results for fitting depth and temperature are compared in Table 4Fb. For 2004 and 2003, the inclusion of depth appears to improve the fit of the exponential variogram model to the data.

All of these analyses were conducted assuming a normal distribution for the observations.

Other distributions such as Poisson or gamma are available in the SPLUS CorrelatedData package or through the use of WinBugs.

***References***

Bannerjee, S., B.P. Carlin, and A.E. Gelfand. 2004. Hierarchical modeling and analysis for Spatial Data, Chapman and Hall/CRC, New York, N.Y.

Cressie, N.A.C. 1993. Statistics for Spatial Data, revised edn, John Wiley & Sons, New York, N.Y.

Pinheiro, J.C., and D.M. Bates. 2000. Mixed-effects models in S and S-PLUS, Sprinter-Verlag, New York.

ajustement, même si l'amélioration de l'ajustement a été minime en 2004 et en 2002 (tableau 4Fa).

On a attiré l'attention durant l'atelier sur la similarité entre le variogramme établi pour la densité de crabes de taille commerciale et celui établi pour les profondeurs observées lors du relevé. Dans la figure 4Fa, on compare les estimations gamma moyennes des classes de distance à l'aide de données datant de 2004 à l'intérieur desquelles la forte tendance linéaire est évidente. Il semble également y avoir une forte tendance similaire pour la température (Fig. 4Fb).

Le logiciel lme() permet d'inclure des covariables comme la profondeur à l'intérieur de la modélisation des variogrammes. On compare au tableau 4Fb des résultats pour l'ajustement de la profondeur et de la température. Pour 2004 et 2003, l'inclusion de la profondeur semble améliorer l'ajustement aux données du modèle de variogramme exponentiel.

On a effectué toutes ces analyses en supposant une distribution normale pour les observations.

On dispose d'autres distributions comme les distributions de Poisson et gamma dans le progiciel SPLUS CorrelatedData ou grâce à l'utilisation de WinBugs.

***Références***

Bannerjee, S., B.P. Carlin, and A.E. Gelfand. 2004. Hierarchical modeling and analysis for Spatial Data, Chapman and Hall/CRC, New York, N.Y.

Cressie, N.A.C. 1993. Statistics for Spatial Data, revised edn, John Wiley & Sons, New York, N.Y.

Pinheiro, J.C., and D.M. Bates. 2000. Mixed-effects models in S and S-PLUS, Sprinter-Verlag, New York.

*Table 4Fa. Comparison of two types of variograms for snow crab commercial density for the 4T annual survey.*

*Tableau 4Fa. Comparaison de deux types de variogrammes pour la densité de crabes des neiges de taille commerciale aux fins du relevé annuel dans 4T*

| Year / Année | Model / Modèle | AIC / Critère d'information d'Akaike |
|---|---|---|
| 2004 | Exponential / Exponentiel | 4422.035 |
| | Spherical / Sphérique | 4424.104 |
| 2003 | Exponential / Exponentiel | 4250.447 |
| | Shperical / Sphérique | 4283.882 |
| 2002 | Expoential / Exponentiel | 4302.790 |
| | Spherical / Sphérique | 4306.983 |

*Table 4Fb. Comparison of exponential variogram models with covariates temperature and depth for snow crab commercial density for the 4T annual survey.*

*Tableau 4Fb. Comparaison de modèles de variogrammes exponentiels avec les covariables température et profondeur pour la densité de crabes des neiges de taille commerciale aux fins du relevé annuel dans 4T*

| Year / Année | Covariate / Covariable | AIC / Critère d'information d'Akaike |
|---|---|---|
| 2004 | None / Aucune | 4422.035 |
| | Depth / Profondeur | 4417.082 |
| | Temperature / Température | 4421.981 |
| 2003 | None / Aucune | 4250.447 |
| | Depth / Profondeur | 4246.955 |
| | Temperature / Température | 4250.298 |
| 2002 | Zone / Aucune | 4302.79 |
| | Depth / Profondeur | 4303.14 |
| | Temperature / Température | 4302.85 |

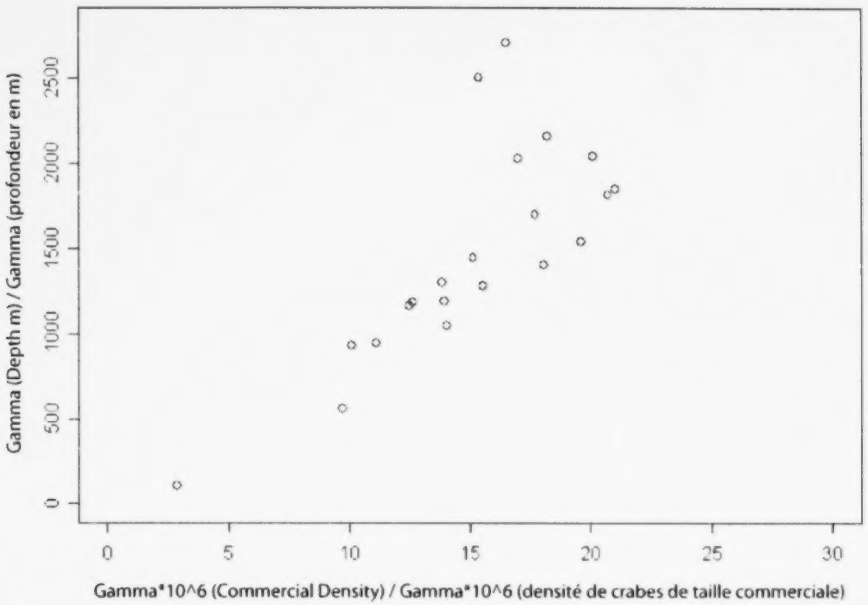



*Figure 4Fa. Comparison of semi-variogram (gamma) estimates for commercial crab density and depth from 2004 4T survey.*

*Figure 4Fa. Comparaison d'estimations (gamma) de semi-variogrammes pour la densité de crabes de taille commerciale et de la profondeur d'après le relevé effectué dans 4T en 2004.*

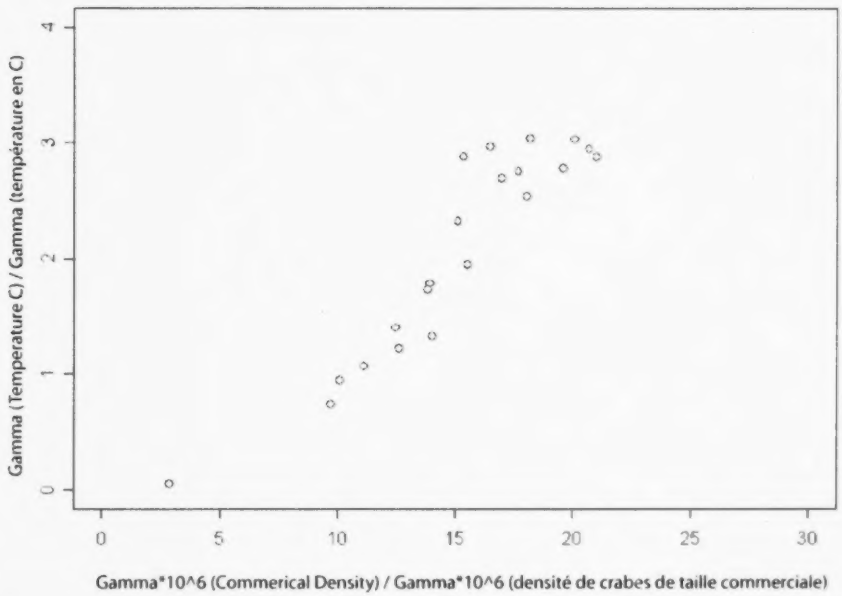



*Figure 4Fb. Comparison of semi-variogram (gamma) estimates for commercial crab density and bottom temperature from 2004 4T survey.*

*Figure 4Fb. Comparaison d'estimations (gamma) de semi-variogrammes pour la densité de crabes de taille commerciale et la température au fond d'après le relevé effectué dans 4T en 2004*

**Appendix 4G. N. Bez. Sensitivity of Kriging Maps and of Estimation Variance to Variogram Model**

The variations of the empirical variograms from one area to the other, or from one year to the next, must be considered relatively to the consequences they have in terms of interpolation maps and in terms of global estimates and estimation variance. In other words: should we be concerned about a 20% difference in the range parameter? In the sill parameter? In the nugget effect?

For illustrative purposes, the maps and the global estimates obtained in Sector 12 in 2004 for four different variogram models were compared. The sill was the same but the range was either 20 or 50 nm and the nugget effect represented either 20% or 50% of the total variance. Figure 4G shows that the interpolation maps were slightly affected by the model. The main difference was in the spatial extension and the level of the hot spots. Larger nugget induced more smoothing. Longer range induced stronger spatial influence of values.

Global estimates ranged from 3891 to 3959 $nb \cdot km^{-2}$ while the arithmetic mean of the data available in Sector 12 was 3426 $nb \cdot km^{-2}$. The four estimates were thus very similar but 16% larger than the arithmetic mean. The kriging procedure affects the weighting given to sample data according to their mutual locations and to their locations in the estimated polygon. In this regards, the numerous samples performed in the Cape Breton subunit participate too strongly in the arithmetic mean while their impact in the kriged estimate was reduced. This illustrates the utility of the geostatistical approach.

Finally, the estimation variance was not sensitive to the model changes that were considered. It ranged from 3.3% or 3.8 % for cases with 20% of nugget effect to 4.6 or 5% for cases with 50% of nugget. This is explained by the fact that there is a strong influence on the polygon estimate by the sample data which are numerous and quasi

**Appendice 4G. N. Bez. Sensibilité des cartes de krigeage et de la variance des estimations au modèle de variogramme**

Il faut examiner les variations des variogrammes empiriques d'une zone à une autre ou d'une année à la suivante par rapport aux conséquences qu'elles ont sur le plan des cartes d'interpolation et des estimations globales et de la variance des estimations. En d'autres mots : Devrions-nous nous inquiéter d'une différence de 20 % à l'intérieur du paramètre zone d'influence? À l'intérieur du paramètre plateau? À l'intérieur du paramètre effet de pépite?

On a comparé à des fins d'illustration les cartes et les estimations globales obtenues dans le secteur 12 en 2004 pour quatre modèles différents de variogrammes. Le plateau était le même, mais la zone d'influence était de 20 ou de 50 NM et l'effet de pépite représentait 20 ou 50 % de la variance totale. La figure 4G montre que le modèle avait une légère influence sur les cartes d'interpolation. La principale différence se situait au niveau de l'extension géographique et des points chauds. Une pépite plus grosse provoquait plus de lissage. Une zone plus longue produisait une influence spatiale plus forte des valeurs.

Les estimations globales allaient de 3 891 à 3 959 $nb \cdot km^{-2}$, tandis que la moyenne arithmétique des données disponibles dans le secteur 12 était de 3 426 $nb \cdot km^{-2}$. Les quatre estimations étaient donc très similaires, mais supérieures de 16 % à la moyenne arithmétique. La procédure de krigeage influence les pondérations attribuées aux données d'échantillon suivant leurs emplacements les unes par rapport aux autres et leurs emplacements également à l'intérieur du polygone estimé. Mentionnons à cet égard que les nombreux échantillons prélevés dans la sous-unité du Cap-Breton occupent une part trop importante à l'intérieur de la moyenne arithmétique, même si l'on a réduit leurs répercussions dans l'estimation krigée, ce qui illustre l'utilité de l'approche géostatistique.

Enfin, la variance des estimations n'était pas sensible aux changements apportés au modèle qu'on a examinés. Elle allait de 3,3 ou de 3,8 % pour les cas où l'effet de pépite atteignait 20 % à 4,6 ou à 5 % pour les cas où cet effet atteignait 50 %. Cela s'explique par le fait que les données d'échantillon, qui sont nombreuses et

regularly spaced. This also shows that the confidence in the abundance estimate over large areas such as sector 12 is large.

quasi régulièrement espacées, ont une forte influence sur l'estimation du polygone. Cela montre aussi que le niveau de confiance dans l'estimation de l'abondance dans de grandes zones comme le secteur 12 est élevé.

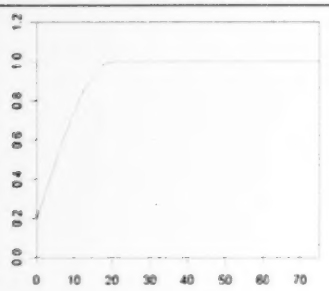
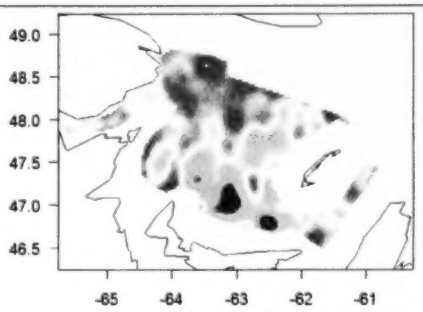
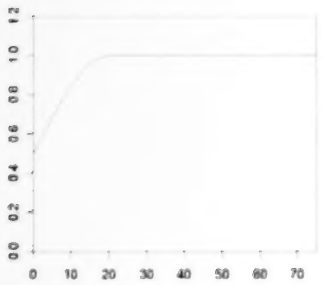
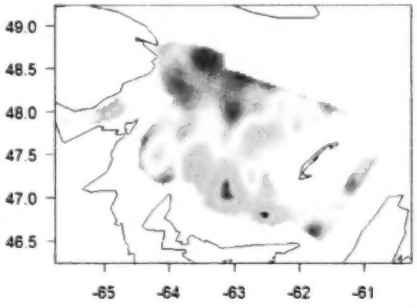
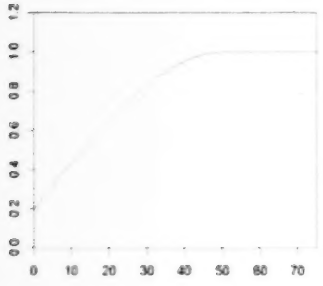
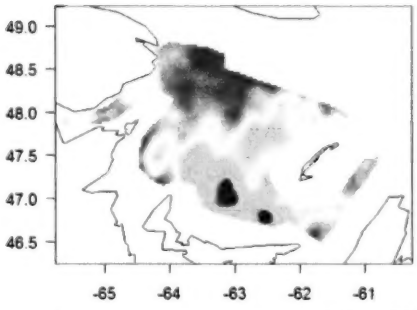
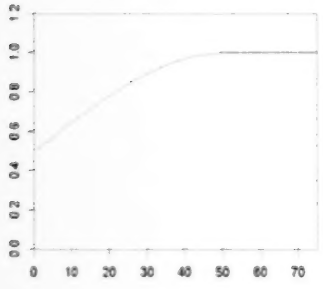
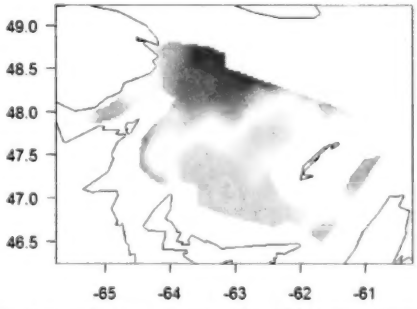
| Model / Modèle | Kriging Map / Carte de krigeage | Polygon Estimation / Estimation du polygone |
|---|---|---|
|  |  | $m^K$ = 3 892<br>CV = 3,8 % |
|  |  | $m^K$ = 3 891<br>CV = 5 % |
|  |  | $m^K$ = 3 949<br>CV = 3,3 % |
|  |  | $m^K$ = 3 959<br>CV = 4,6 % |

*Fig 4G. Impact of model parameter on the kriging interpolation and on the polygon global estimation and estimation variance. Sector 12, 2004, all data in the polygon. Arithmetic mean = 3426 nb·km$^{-2}$, sample variance = 12769866 nb$^2$·km$^{-4}$. All maps have the same colour scale.*

*Fig. 4G. Répercussions du paramètre du modèle sur l'interpolation du krigeage et sur l'estimation globale et la variance de l'estimation du polygone. Secteur 12, 2004; toutes les données sont incluses dans le polygone. La moyenne arithmétique = 3 426 nb·km$^{-2}$ et la variance de l'échantillon = 12 769 866 nb$^2$·km$^{-4}$. On a utilisé la même gamme de couleurs pour toutes les cartes.*

**Appendix 4H. D. Marcotte. Cross-Validation**

Cross-validation is a technique based on the idea of re-estimating each known data point by using its neighbours. Comparing the estimated data values with the observed values, errors are defined for each data point. Different statistics can be computed on the errors to evaluate the quality of the results with the tested procedure. By changing the procedure, one can re-compute error statistics and decide upon which procedure is better.

Typically, cross-validation could be used to:

- Choose between two different variogram models (e.g. one with a small nugget effect and a short range vs one with a larger nugget and larger range; spherical vs exponential; etc.)
- Iterate over the whole model parameter space to find the parameter set giving the smallest error statistic
- Compare the efficiency of different kriging neighbourhoods
- Compare the efficiency of ordinary kriging vs kriging with a drift or other types of kriging
- Check that the current model is not over-estimating or under-estimating the precision of kriging estimates
- Compare efficiency of kriging vs. other estimation methods
- Check the statistical homogeneity of the studied area by looking at spatial displays of signs and amplitudes of cross-validation errors

Defining $e_i = Z_i - Z_i^*$ (raw cross-validation error) and $n_i = \frac{e_i}{\sigma_{ki}}$ (normalized cross-validation error) where $Z_i$ is the observed data at point "i", $Z_i^*$ is the

**Appendice 4H. D. Marcotte. Contre-validation**

La contre-validation est une technique fondée sur l'idée de réestimer chaque point de données connu en utilisant les points voisins. En comparant les valeurs de données estimées aux valeurs observées, on définit des erreurs pour chaque point de données. On peut calculer différentes statistiques en se fondant sur les erreurs pour évaluer la qualité des résultats avec la procédure mise à l'essai. En modifiant la procédure, il est possible de recalculer des statistiques d'erreur et de déterminer la meilleure procédure.

On pouvait habituellement utiliser la contre-validation pour :

- Choisir entre deux modèles de variogrammes différents (p. ex. un avec un petit effet de pépite et une courte zone d'influence par opposition à un avec un effet de pépite plus élevé et une zone d'influence plus vaste, un modèle sphérique plutôt qu'exponentiel, etc.).
- Appliquer une itération à tout l'espace des paramètres du modèle afin de trouver l'ensemble de paramètres donnant la statistique d'erreur la plus petite.
- Comparer l'efficacité de différents voisinages de krigeage.
- Comparer l'efficacité du KO par opposition au KDE ou d'autres types de krigeage.
- Vérifier que le modèle actuel ne surestime pas ou ne sous-estime pas la précision des estimations par krigeage.
- Comparer l'efficacité du krigeage à celle d'autres méthodes d'estimation.
- Vérifier l'homogénéité statistique de la zone étudiée en examinant les affichages spatiaux de signes et les amplitudes d'erreurs de contre-validation.

Si l'on établit que $e_i = Z_i - Z_i^*$ (l'erreur de contre-validation brute) et que $n_i = \frac{e_i}{\sigma_{ki}}$ (l'erreur de contre-validation normalisée), où $Z_i$ est la

estimated value by kriging with its neighbours the point "i" and $\sigma_{ki}$ the corresponding kriging standard deviation, the following different statistics could be used:

a. Average value of the errors or normalised errors $\frac{1}{n}\sum_i e_i$ or $\frac{1}{n}\sum_i n_i$ should be close to zero (absence of bias)

b. Average values of absolute errors or squared errors should be as small as possible : $\frac{1}{n}\sum_i |e_i|$ or $\frac{1}{n}\sum_i e_i^2$ minimum; the former statistic (absolute) being more robust to a few extreme values;

c. Square root of mean square normalized errors should be close to one: $\left(\frac{1}{n}\sum_i n_i^2\right)^{0.5} \approx 1$; a value larger than one means the current model is predicting better precision (smaller kriging standard deviation) than really achieved; the opposite being true for a value lower than one. A discrepancy of say 10% is easily tolerated, sometimes more. If the discrepancy is too large, either homogeneous zones should be redefined or the model used in kriging re-evaluated;

***Application to 4T Snow Crab***

The cross-validation technique was used at the meeting to compare two alternatives fitted to an experimental variogram.

donnée observée au point « i », $Z_i^*$ est la valeur estimée par krigeage avec ses voisins le point « i » et $\sigma_{ki}$, l'écart-type de krigeage correspondant, on pourrait utiliser les différentes statistiques suivantes :

a. La valeur moyenne des erreurs ou des erreurs normalisées $\frac{1}{n}\sum_i e_i$ ou $\frac{1}{n}\sum_i n_i$ devrait approcher zéro (absence de biais);

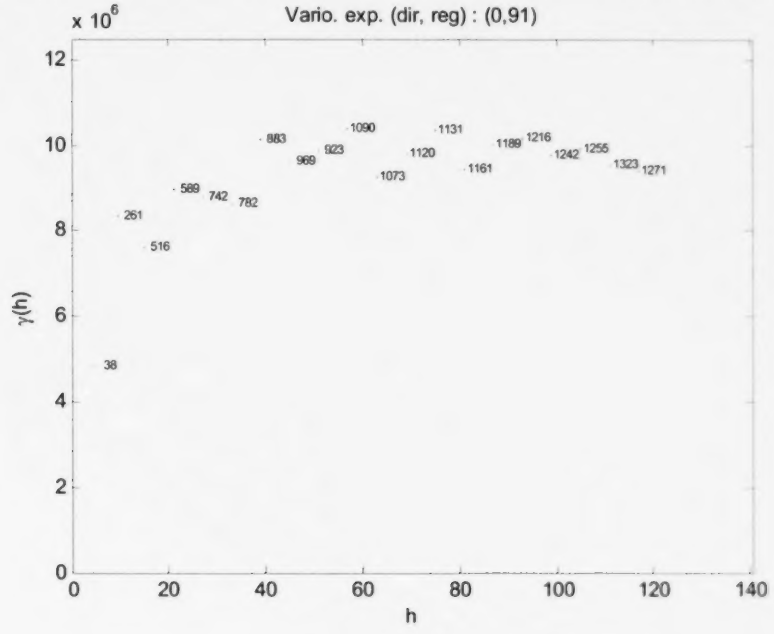
b. Les valeurs moyennes d'erreurs absolues ou d'erreurs quadratiques devraient être aussi peu élevées que possible : $\frac{1}{n}\sum_i |e_i|$ ou $\frac{1}{n}\sum_i e_i^2$ minimum; la première statistique (celle des erreurs absolues) est plus robuste par rapport à quelques valeurs extrêmes;

c. La racine carrée des erreurs quadratiques moyennes normalisées devrait se rapprocher de un : $\left(\frac{1}{n}\sum_i n_i^2\right)^{0.5} \approx 1$; une valeur supérieure à un signifie que le modèle actuel prédit plus précisément (écart-type de krigeage plus petit) que ce qu'on atteint réellement, le contraire étant vrai pour une valeur inférieure à un. On tolère facilement une divergence de 10 % disons, et parfois de plus de 10 %. Si la divergence est trop grande, on devrait redéfinir des zones homogènes ou réévaluer le modèle évalué dans le cadre du krigeage.

***Application au crabe des neiges de 4T***

On a utilisé la technique de contre-validation à la réunion pour comparer deux solutions de rechange ajustées à un variogramme expérimental.

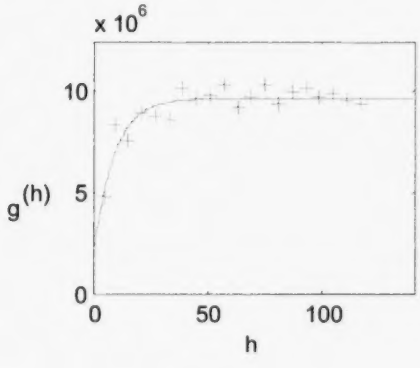
2003 : Data from areas 12+E+F CBsubunit / 2003 : Données des zones 12, E et F de la sous-unité du C-B




On this experimental variogram, two alternative models were compared, one with a small nugget, small range, ant the other with a large nugget, large range. Visually, the fits are almost equivalent.

a) The small nugget - small range exponential model:

Vario. Exp. (dir, dip, vreg) : (0,0,91)




Exponential / Exponentiel
C0 = 1.7e06
C = 7.9e06
a = 26 km

Mod. Iso. : type a, c
1. 1 $1.744e^{+006}$
2. 8.783 $7.93e^{+006}$

1 – small / pépite
2 – exponential / exponentiel

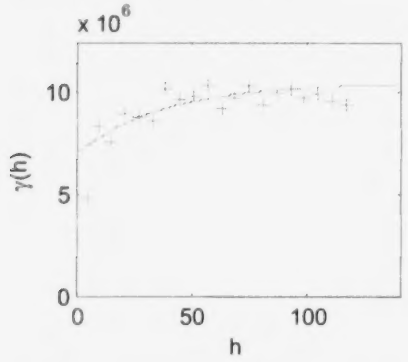
Sur ce variogramme expérimental, on a comparé deux modèles de rechange, un avec un petit effet de pépite et une petite zone d'influence et l'autre avec un gros effet de pépite et une vaste zone d'influence. Visuellement, les ajustements sont presque équivalents.

a) Le modèle exponentiel avec un petit effet de pépite et une petite zone d'influence :

b) The large nugget – large range exponential model:

b) Le modèle exponentiel avec un gros effet de pépite et une vaste zone d'influence :

Vario. Exp. (dir, dip, vreg) : (0,0,91)




Exponential / Exponentiel
C0 = 7e06
C = 3.5e06
a = 115 km

Mod. Iso. : type a, c
1. 1 $7e^{+006}$
2. 38.63 $3.5e^{+006}$

1 – small / pépite
2 – exponential / exponentiel

The cross-validation statistics obtained are:

Les statistiques de contre-validation obtenues sont les suivantes :

| Statistic / Statistique | Model A / Modèle A | Model B / Modèle B |
|---|---|---|
| $\frac{1}{n}\sum_{i=1}^{n}\lvert Z_i - Z_i^*\rvert$ | 1979 | 2032 |
| $\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Z_i - Z_i^*}{\sigma_k}\right)^2\right)^{0.5}$ | 0.93 | 0.92 |

A 2.6% improvement on the absolute error with model A (small nugget-short range) compared to model B was observed. Thus, the model A provides a better cross-validation than model B. Both gave acceptable normalized error standard deviation, suggesting that both models predict correctly the precision of the estimates they provide.

On a observé une amélioration de 2,6 % pour l'erreur absolue avec le modèle A (petit effet de pépite et petite zone d'influence) comparativement au modèle B. Le modèle A fournit donc une meilleure contre-validation que le modèle B. Les deux modèles donnaient un écart-type d'erreur normalisée acceptable, ce qui laisse supposer qu'ils prédisent correctement la précision des estimations qu'ils fournissent.

The cross-validation was also used to compare results of ordinary kriging (OK) to results of kriging with an external drift (KED) for sector 1 (Baie des Chaleurs) with and without pseudo-zeros. Results are given in the next table. They indicate that:

On a aussi utilisé la contre-validation pour comparer les résultats du KO aux résultats du KDE pour le secteur 1 (la baie des Chaleurs) avec et sans pseudozéros. Les résultats figurent au tableau qui suit. Ils indiquent que :

- KED provides a (1110-1222) / 1222=9.2% improvement of the cross-validation statistic (case without pseudo-zeros)
- OK with pseudo-zeros shows a strong bias (although this bias is over-estimated by this test due to the lack of data at smaller depths); in all other cases, there is no significant bias
- KED is almost insensitive to the presence of pseudo-zeros. Again this conclusion is partly due to the fact that very few data were taken at smaller depths, say between 20m and 60m.

- Le KDE assure une amélioration de 1 110-1 222) / 1 222 = 9,2 % de la statistique de contre-validation (cas sans pseudozéros);
- Le KO avec pseudozéros indique un biais élevé (même si cet essai surestime ce biais en raison du manque de données à des profondeurs inférieures); il n'y a dans aucun cas de biais significatif;
- Le KDE est presque insensible à la présence de pseudozéros. Cette conclusion est encore une fois en partie attribuable au fait qu'on a prélevé très peu de données à des profondeurs inférieures, entre 20 et 60 m disons.

| **Ordinary kriging (OK) and kriging with an external drift (KED) / Krigeage ordinaire et krigeage avec dérive externe** | **Without pseudo-zeros / Sans pseudozéros** | | **With pseudo-zeros / Avec pseudozéros** | |
|---|---|---|---|---|
| | **OK / KO** | **KED / KDE** | **OK / KO** | **KED / KDE** |
| # crabes in polygon (x $10^6$) / Nombre de crabes dans le polygone (x $10^6$) | 8.35 | 7.33 | 6.25 | 7.33 |
| Std of estimation (x $10^6$) / norme de l'estimation (x $10^6$) | 1.41 | 1.52 | 1.22 | 1.25 |
| Cross-validation / contre-validation $\frac{1}{n}\sum_{i=1}^{n}\lvert Z_i - Z_i^*\rvert$ | 1222 | 1110 | 1354 | 1106 |
| Cross-validation / contre-validation $\frac{1}{n}\sum_i e_i$ | 46 | 62 | 832 | 40 |

Thus the cross-validation indicates KED should provide better estimates than OK for this area.

La contre-validation indique donc que le KDE devrait fournir pour cette zone de meilleures estimations que le KO.

**Appendix 4I. D. Marcotte. Kriging with External Drift**

It is supposed that the observed variable Z(x) possesses the following decomposition:

$$Z(x) = m(x) + W(x)$$

where W(x) is a second-order stationary process with mean 0 and m(x) is a drift function giving E[Z(x)]. This drift can be deterministic and function uniquely of the "x" coordinates (case of universal kriging or kriging with drift). Alternately, the drift function can be a linear function of a secondary or auxiliary variable, i.e. $m(x) = a_0 + a_1 Y(x)$. Here for instance, we can assume it is the case with Y(x) being the depth variable. In that case, we speak of kriging with an external drift.

The kriging system with external drift is:

$$\begin{bmatrix} K & F \\ F' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} k_s \\ f \end{bmatrix}$$

where

K is the nxn data covariances matrix;
F is n x nc : nc (=2) drift functions at data points $x_i$

f is nc x 1 : nc drift functions at estimation point $x_0$

$\mu$ is the nc x 1 vector of Lagrange multipliers

If the drift is assumed to be related to depth "d", then, we will get:

$$F = \begin{bmatrix} 1 & d_1 \\ 1 & d_2 \\ \bullet & \bullet \\ 1 & d_n \end{bmatrix}, \quad f = \begin{bmatrix} 1 \\ d_0 \end{bmatrix}$$

the index "0" refers to the estimation point.

**Appendice 4I. D. Marcotte. Krigeage avec dérive externe**

On suppose que la variable observée Z(x) est décomposée de la façon suivante :

$$Z(x) = m(x) + W(x),$$

où W(x) est un processus stationnaire du 2e ordre dont la moyenne est 0 et m(x) est une fonction de dérive donnant E[Z(x)]. Cette dérive peut être déterministe et fonction uniquement des coordonnées « x » (le cas du krigeage ou du krigeage avec dérive). Autrement, la fonction de dérive peut être une fonction linéaire d'une variable secondaire ou auxiliaire, c'est-à-dire $m(x) = a_0 + a_1 Y(x)$. Ici, par exemple, nous pouvons supposer que c'est le cas, Y(x) étant la variable de profondeur. Dans ce cas, nous parlons de KDE.

Le système de KDE est le suivant :

$$\begin{bmatrix} K & F \\ F' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} k_s \\ f \end{bmatrix},$$

où

K est la matrice des covariances de données nxn;
F est n x nc : nc (=2) fonctions de dérive aux points de données $x_i$;
f est nc x 1 : nc fonctions de dérive au point d'estimation $x_0$;
$\mu$ est le vecteur nc x 1 des multiplicateurs de Lagrange.

Si l'on suppose que la dérive est reliée à la profondeur « d », nous obtiendrons alors :

$$F = \begin{bmatrix} 1 & d_1 \\ 1 & d_2 \\ \bullet & \bullet \\ 1 & d_n \end{bmatrix}, \quad f = \begin{bmatrix} 1 \\ d_0 \end{bmatrix},$$

où l'indice « 0 » désigne le point d'estimation.

The coefficients $a_0$ and $a_1$ need not be known "a priori" to use kriging with external drift. The last two equations of the kriging system, i.e. $F\lambda = f$, ensure the estimate will be unbiased whatever the values of $a_0$ and $a_1$. In fact, what happens is that these coefficients are implicitly estimated when solving the kriging system. This means that for each different neighbourhood (e.g. with moving neighbourhood), the coefficients will be re-estimated implicitly allowing for great flexibility in the exact shape of the drift function.

Il n'est pas nécessaire de connaître « a priori » les coefficients $a_0$ et $a_1$ pour utiliser le KDE. Les deux dernières équations du système de krigeage, c'est-à-dire $F\lambda = f$, fournissent l'assurance que l'estimation ne sera pas biaisée quelles que soient les valeurs de $a_0$ et $a_1$. En fait, ce qui arrive, c'est que ces coefficients sont implicitement estimés lorsqu'on solutionne le système de krigeage. Cela signifie que pour chaque voisinage différent (comme un voisinage mobile), les coefficients seront réestimés implicitement, ce qui allouera une plus grande souplesse à l'intérieur de la forme exacte de la fonction de dérive.

The K matrix contains the nxn covariances between data points. These covariances are the covariances of W(x), i.e. of the residuals. In some cases, this covariance can be approximated by the Z(x) covariances (or variogram). In other cases (when the slope $a_1$ is strong) it will be necessary to find other ways to estimate the covariance model. Among the many methods available, cross-validation can be helpful. In the snow-crab case, because the correlation between depth and density is weak, the coefficient $a_1$ is small and the variogram computed with Z(x) can be used as a safe approximation for the variogram of W(x).

La matrice K renferme les covariances nxn entre les points de données. Ces covariances sont les covariances de W(x), c'est-à-dire des résidus. Dans certains cas, on peut lisser une telle covariance au moyen des covariances Z(x) (ou du variogramme). Dans d'autres (lorsque la pente $a_1$ est forte), il faudra trouver d'autres moyens d'estimer le modèle de covariances. Mentionnons que parmi les nombreuses méthodes disponibles, la contre-validation peut être utile. Dans le cas du crabe des neiges, parce que la corrélation entre la profondeur et la densité est faible, le coefficient $a_1$ est petit et le variogramme calculé avec Z(x) peut servir d'approximation sûre pour le variogramme de W(x).

When using KED, because a model of drift is implicitly fitted to the data, the kriging neighbourhood must be large enough (i.e. with enough data points) to ensure a stable estimate of the drift and to avoid possible degenerate configurations that would lead to a singular kriging matrix (ex. if the n points used in the kriging neighbourhood have exactly the same depth, then the second column of matrix F is proportional to the first one and the kriging matrix is singular).

Lorsqu'on utilise le KDE, parce qu'un modèle de dérive est implicitement ajusté aux données, le voisinage de krigeage doit être suffisamment large (c'est-à-dire avoir assez de points de données) pour assurer une estimation stable de la dérive et pour éviter d'éventuelles configurations dégénérées qui mèneraient à une matrice de krigeage singulière (par exemple, si les points n utilisés dans le voisinage de krigeage ont exactement la même profondeur, la 2$^e$ colonne de la matrice F est alors proportionnelle à la première et la matrice de krigeage est singulière).

Note that KED assumes a linear relationship between E[Z(x)] and depth. This hypothesis has more chances to be fulfilled with local neighbourhoods than with a global one. In this last case, the very large depth could have the same expectation than the very shallow depth (0) thus, the link is not linear. In that case, KED will more or less boils down to OK. One alternative is to transform the depth variable in such a way that the relationship with E[Z(x)] becomes linear (ex. taking |80m-d| instead of "d" could help linearize the relationship over the whole range of depth). Again,

Noter qu'on suppose dans le cadre du KDE une relation linéaire entre E[Z(x)] et la profondeur. Cette hypothèse est plus susceptible de se réaliser avec des voisinages locaux qu'avec un voisinage global. Dans ce dernier cas, la très grande profondeur pourrait avoir la même valeur probable que la très petite profondeur (0); le lien n'est donc pas linéaire. Dans ce cas, le KDE reviendra plus ou moins à du KO. Une des solutions de rechange consiste à transformer la variable de profondeur de façon à ce que la relation avec E[Z(x)] devienne linéaire (par

cross-validation could be useful to choose between different possible transformations of the auxiliary variable (note that the primary variable is not transformed to avoid problems in the back transformation).

Note that when using KED, the auxiliary variable must be known at all data points and at all estimation points. If it is not the case, it must first be estimated to fill the grid. Normally, the auxiliary variable is less varying (i.e. smoother) and has better spatial continuity than the variable of interest.

***Example of KED in Sector 1 with depth as external drift***

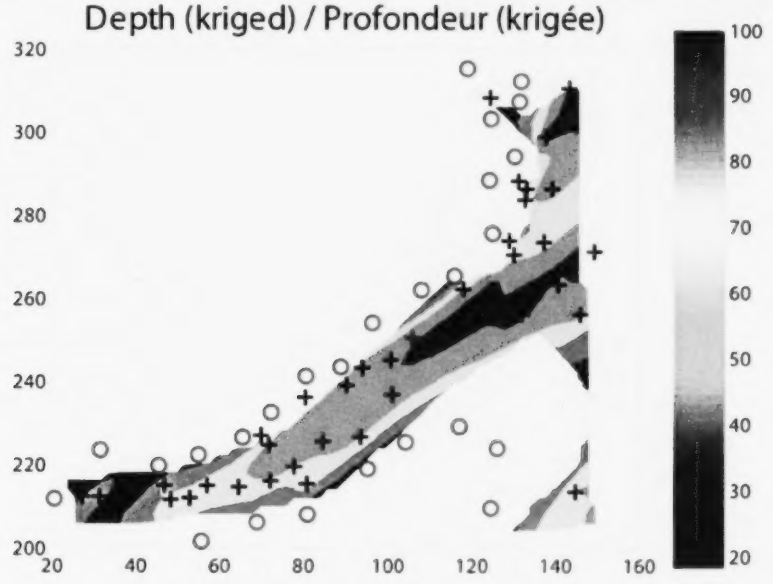
First, the depth is estimated everywhere:

exemple, le fait de retenir |80m-d| plutôt que « d » pourrait aider à linéariser la relation à l'intérieur de toute la gamme des profondeurs). Encore une fois, la contre-validation pourrait être utile pour choisir entre différentes transformations possibles de la variable auxiliaire (noter qu'on ne transforme pas la variable primaire pour éviter des problèmes dans le cadre de la retransformation).

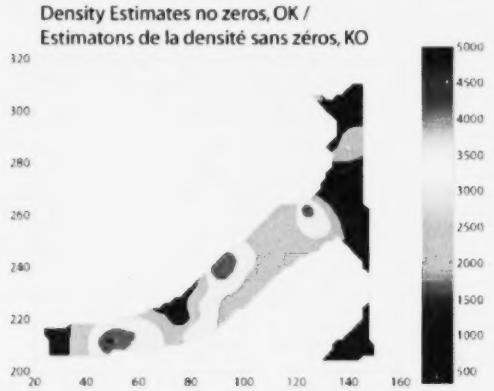
Noter également que quand on utilise le KDE il faut connaître la variable auxiliaire à tous les points de données et à tous les points d'estimation. Si ce n'est pas le cas, on doit premièrement l'estimer pour remplir la grille. Normalement, la variable auxiliaire varie moins (c'est-à-dire est plus lisse) et a une meilleure continuité spatiale que la variable d'intérêt.

***Exemple de KDE dans le secteur 1 faisant appel à la profondeur comme dérive externe***

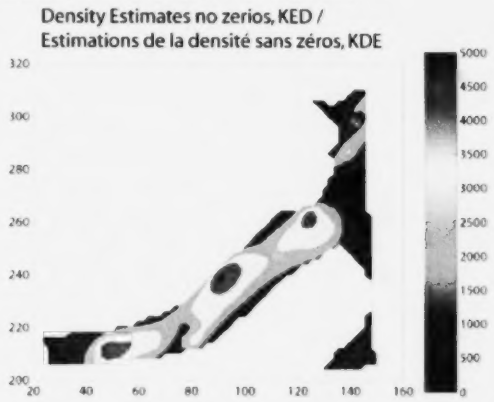
On estime premièrement partout la profondeur :



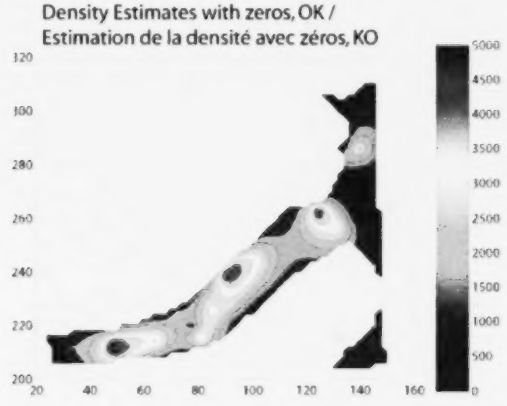
# OK No zeros / KO sans zéros



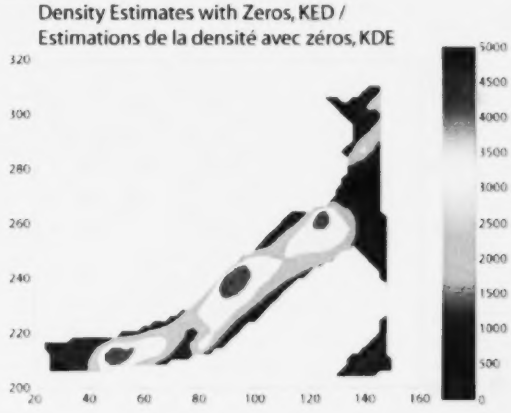

# KED No zeros / KDE sans zéros




# OK With zeros / KO avec zéros




# KED With zeros / KDE avec zéros




Note how the KED maps with and without pseudo-zeros are almost identical. This is due to the fact that the interpolated points were mostly at depths between in the [40-120] m range. More important is the better alignment with the channel orientation of KED compared to KO due to the effect of using depth as external drift. Cross-validation (Appendix 4H) indicates also that KED estimates are expected to be more precise than OK estimates due to the additional relevant information provided by depth.

Noter aussi en quoi les cartes de KDE avec et sans pseudozéros sont presque identiques. Cela tient au fait que les points interpolés se trouvaient en majorité à des profondeurs qui allaient de 40 à 120 m. Le meilleur alignement avec l'orientation du chenal de KDE comparativement au KO attribuable au fait qu'on utilise la profondeur comme dérive externe est plus important. La contre-validation (appendice 4H) indique également que les estimations par KDE devraient être plus précises que celles établies par KO en raison des données pertinentes additionnelles que fournit la profondeur.

**Appendix 4J. N. Bez. Taking into Account the Relationship with an Environmental Parameter**

The relationship between the studied variable and external forcing parameters can lead to several but non equivalent approaches. When there exists a linear relationship between the mean of the studied variable and that of environmental parameters, kriging with external drift (KED) represents a practical compromise for the incorporation of auxiliary variable in the estimation of snow crab density and abundance.

The general principle is to perform an ordinary kriging (OK) with the additional constraint to filter out, at the neighbourhood scale, the linear influence of the external parameter. So doing, if everything is unchanged in the kriging process (same variogram model, same neighbourhood, same sample points), the estimation variance obtained with KED cannot be smaller than the one obtained with OK. As a matter of fact, OK weights are those that produce the smallest estimation of variance by definition. Those of KED are thus sub-optimal.

For Area 12, the use of bathymetry considered as a proxy for the crab habitat, has been investigated. Bathymetry is known at sampling points and has been interpolated on Area 12 fishing ground (Fig. 4J). The global relationship between crab density and bathymetry is not linear but rather bell-shaped with a dome at 40 fathoms. We thus used the following simple transformation of the bathymetry to turn the relationship into a linear one:

bathymetry → abs(40-bathymetry)

The use of this transformed bathymetry as external drift led to the interpolated map shown in Figure 4J. The map of the absolute difference between OK and KED show the areas where the use of an external drift impacts most the results. Sharp changes occur in short distances which should be investigated further.

The use of such external drift is believed to be a good alternative to the use of pseudo zeroes as it naturally leads the crab density to decrease at the

**Appendice 4J. N. Bez. Prise en compte de la relation avec un paramètre environnemental**

La relation entre la variable étudiée et des paramètres de contrainte externes peuvent mener à plusieurs approches, mais qui ne sont pas équivalentes. Lorsqu'il existe une relation linéaire entre la moyenne de la variable étudiée et celle de paramètres environnementaux, le KDE représente un compromis pratique pour l'intégration d'une variable auxiliaire à l'estimation de la densité et de l'abondance du crabe des neiges.

Le principe général consiste à effectuer un KO avec la contrainte additionnelle pour éliminer, à l'échelle du voisinage, l'influence linéaire du paramètre externe. En agissant ainsi, si rien ne change à l'intérieur du processus de krigeage (si le modèle de variogramme, le voisinage et les points d'échantillon restent les mêmes), la variance de l'estimation obtenue avec le KDE ne peut être plus petite que celle obtenue avec le KO. De fait, les pondérations du KO sont celles qui produisent l'estimation la plus petite de la variance par définition. Celles du KDE sont donc sous-optimales.

Dans le cas de la zone 12, on a étudié l'utilisation de la bathymétrie considérée comme une approximation de l'habitat du crabe. On connaît la bathymétrie aux points d'échantillonnage et on l'a interpolée au-dessus du lieu de pêche de la zone 12 (Fig. 4J). La relation globale entre la densité de crabes et la bathymétrie n'est pas linéaire; elle a plutôt la forme d'une cloche dont le dôme se situe à 40 brasses. Nous avons donc utilisé la transformation simple qui suit de la bathymétrie pour convertir la relation en une relation linéaire :

bathymétrie → abs(40-bathymétrie)

L'utilisation de cette bathymétrie transformée sous forme de dérive externe a mené à la carte interpolée illustrée à la figure 4J. La carte de la différence absolue entre le KO et le KDE montre les zones où l'utilisation d'une dérive externe influence le plus les résultats. Il se produit des changements de forme sur de courtes distances qu'il faudrait étudier davantage.

On croit que l'utilisation d'une telle dérive externe est une bonne solution de rechange à l'emploi de pseudozéros, étant donné qu'elle entraîne

edge of its distribution.

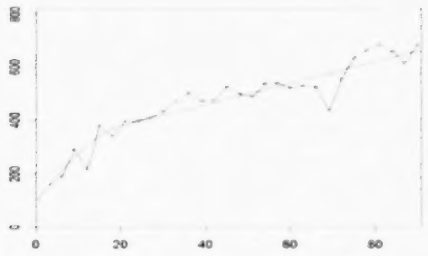
Compared to a more sophisticated approach like co-kriging, kriging with external drift requires knowledge of the external variable at any single sampling point and also at any single point where the interpolation is made. This often requires interpolating the external variable prior to the computations but this may introduce problems if the nugget effect of the variogram model of the external variable gets a strong nugget effect (strong discrepancies of the external variable between sampling points and grid points).

Research and tests are required before using this approach in assessments. In addition to the previous remarks, the impact of the characteristics of the neighbourhood search on (i) the kriging map and (ii) the global estimate must be clearly investigated and controlled. In addition, this approach should ideally be based on the variogram of the residuals which is not accessible for the particular reason that one of the strength of the method is to avoid having to compute residuals. Ways to handle this kind of question exist and should be developed (cross validation for instance).

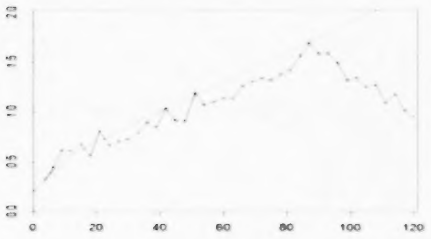
naturellement une diminution de la densité de crabes en bordure de la zone de distribution de ces derniers.

Comparativement à une méthode plus sophistiquée comme le co-krigeage, il faut pour le KDE connaître la variable externe à tout point d'échantillonnage et également à tout point où l'interpolation s'effectue. Cela exige souvent l'interpolation de la variable externe avant les calculs, ce qui risque cependant de soulever des problèmes si l'effet de pépite du modèle de variogramme de la variable externe entraîne un effet de pépite marqué (de fortes divergences de la variable externe entre les points d'échantillonnage et les points de grille).

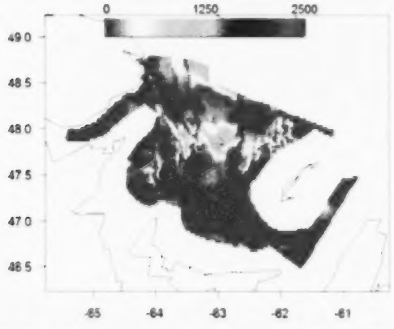
Il faut effectuer des recherches et des essais avant d'utiliser cette méthode à l'intérieur d'évaluations. Notons, en plus des remarques précédentes, qu'on doit clairement étudier et contrôler les répercussions des caractéristiques de la recherche dans le voisinage sur i) la carte de krigeage et ii) l'estimation globale. En plus, cette méthode devrait également reposer sur le variogramme des résidus qui est inaccessible pour la raison particulière que l'une des forces de la méthode est le fait qu'elle évite d'avoir à calculer des résidus. Il existe des moyens, qu'on devrait mettre au point, de traiter ce genre de question (la contre-validation, par exemple).



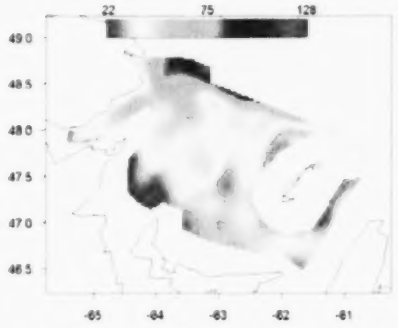
Variogram (empirical and model) of Depth / Variogramme (empirique et modèle) de la profondeur



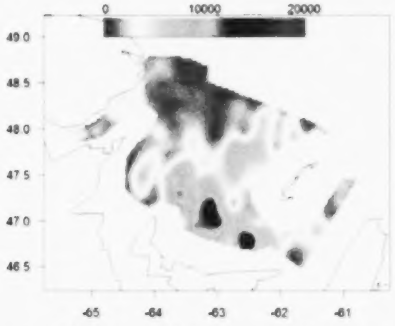
Variogram (empirical and model) of Commerical Densities / Variogramme (empirique et modèle) des densités de crabes de taille commerciale



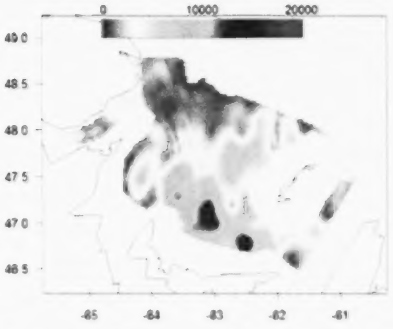
Absolute Difference Between the Ordinary Kriging and the Kriging with External Drift / Différence absolue entre le krigeage ordinaire et le krigeage avec dérive externe



Ordinary Kriging of Depth / Krigeage ordinaire de la profondeur



Ordinary Kriging of Commercial Density / Krigeage ordinaire de la densité de crabes de taille commerciale



Kriging of Commercial Density with Depth as External Drift / Krigeage de la densité de crabes de taille commerciale à l'aide de la profondeur comme dérive externe

*Figure 4J. Kriging with bathymetry as external drift.*

*Figure 4J. Krigeage à l'aide de la bathymétrie comme dérive externe*

**Appendix 4K. N. Bez. Numerical Approximation for the Computation of the Geostatistical Estimation Variance**

The geostatistical estimation of the crab density over a polygon requires the computation of the integration of the variogram values over all possible distances of two points taken at random in the polygon. Except for simple polygons, this has no analytical solution and one must use a numerical solution in which the key step is the discretization of the polygon of interest (the smaller the discretization, the better the approximation). A compromise must then be found between the computer memory capacities (increases with smaller discretization) and the size of the discretization.

In Figure 4K, one can see that, for Area 12 in 2004, the coefficient of variation decreases as we move from a 50 to a 2 nautical miles discretization. Using a 1 n.mi. by 1 n.mi. grid multiplies the computation time by 4 and does not improve the result which already reached the asymptote for a 2 by 2 n.mi. grid.

**Appendice 4K. N. Bez. Approximation numérique pour le calcul de la variance de l'estimation géostatistique**

Il faut pour estimer géostatistiquement la densité de crabes au-dessus d'un polygone calculer l'intégration des valeurs du variogramme pour toutes les distances possibles entre deux points choisis au hasard à l'intérieur de ce polygone. Sauf dans le cas des polygones simples, il n'y a pas pour ce faire de solution analytique et on doit utiliser une solution numérique à l'intérieur de laquelle la principale étape ou opération est la discrétisation du polygone d'intérêt (plus la discrétisation est limitée, plus l'approximation est de qualité). On doit donc alors trouver un compromis entre la capacité de mémorisation de l'ordinateur (qui augmente quand la discrétisation est limitée) et l'ampleur de cette dernière.

On peut constater à la figure 4K qu'en 2004 pour la zone 12 le coefficient de variation diminue à mesure que nous passons d'une discrétisation de 50 à 2 milles marins. L'utilisation d'une grille de 1 NM x 1 NM quadruple le temps de calcul et n'améliore pas le résultat qui a déjà atteint l'asymptote pour une grille de 2 x 2 NM.

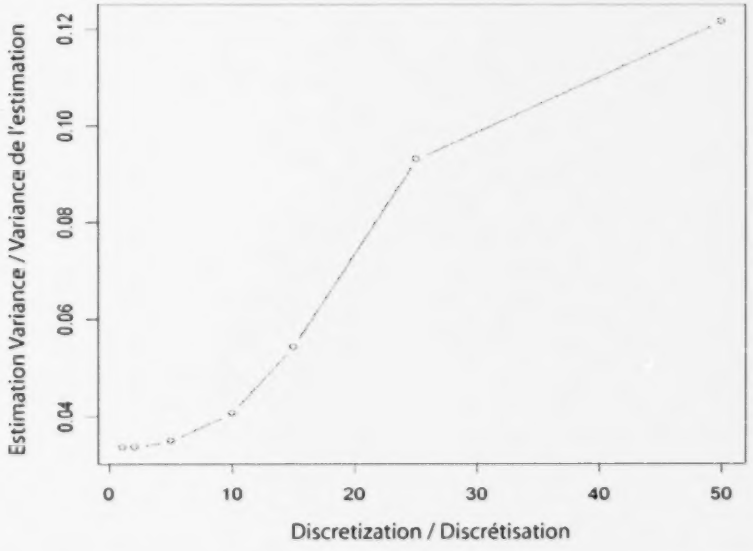



*Figure 4K. Evolution of the coefficient of variation for a smaller and smaller discretization parameter for Area 12, 2004, commercial density.*

*Figure 4K. Évolution du coefficient de variation pour un paramètre de discrétisation de plus en plus petit aux fins du calcul de la densité de crabes de taille commerciale en 2004 dans la zone 12*